

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
10 May 2001 (10.05.2001)

PCT

(10) International Publication Number  
**WO 01/32918 A1**

- (51) International Patent Classification<sup>7</sup>: C12Q 1/48, West Chester, PA 19382 (US). LEE, John, C.; 1706A Handball Lane, Indianapolis, IN 46260 (US).  
C12N 9/12, 15/54
- (21) International Application Number: PCT/US00/30542
- (22) International Filing Date: 6 November 2000 (06.11.2000)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
60/163,901 5 November 1999 (05.11.1999) US  
09/706,426 3 November 2000 (03.11.2000) US
- (71) Applicants: SMITHKLINE BEECHAM CORPORATION [US/US]; One Franklin Plaza, Philadelphia, PA 19130 (US). SMITHKLINE BEECHAM PLC. [GB/GB]; Two New Horizons Court, Brentford, Middlesex TW8 9EP (GB).
- (72) Inventors: KASSIS, Shouki; 7 Lee Road, Audubon, PA 19403 (US). ANNAN, Roland, S.; 1008 Radley Drive,
- (74) Agents: ANDERSEN, Robert, L. et al.; Ratner & Prestia, 301 One Westlakes (Berwyn), P.O. Box 980, Valley Forge, PA 19482-0980 (US).
- (81) Designated State (*national*): JP.
- (84) Designated States (*regional*): European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).
- Published:  
— With international search report.  
— Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: SPECIFIC PEPTIDE SUBSTRATE FOR HUMAN YAK1, YAK3A, AND YAK3B AND YEAST YAK1 PROTEIN KINASES

(57) Abstract: hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptides and polynucleotides, methods for producing such polypeptides by recombinant techniques, and methods of using Ser-164 and these polypeptides and polynucleotides to determine the effect on phosphorylation are disclosed. Also disclosed are methods for utilizing hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptides and polynucleotides in the design of protocols for the treatment of bone loss including osteoporosis; inflammatory diseases such as Adult Respiratory Disease Syndrome (ARDS), Rheumatoid arthritis, Osteoarthritis, Inflammatory Bowel Disease (IBD), psoriasis, dermatitis, asthma, allergies; infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; HIV-associated cachexia and other immunodeficiency disorders; septic shock; pain; injury; cancers; anorexia; bulimia; Parkinson's disease; cardiovascular disease including restenosis, atherosclerosis, acute heart failure, myocardial infarction; hypotension; hypertension; urinary retention; angina pectoris; ulcers; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome, among others, and diagnostic assays for such conditions.

WO 01/32918 A1

- 1 -

## SPECIFIC PEPTIDE SUBSTRATE FOR HUMAN YAK1, YAK3A, AND YAK3B AND YEAST YAK1 PROTEIN KINASES

### Related Applications

This application claims priority of U.S. Provisional Application Serial No. 60/163,901,  
5 filed on November 5, 1999.

### Field of the Invention

This invention relates to a newly identified peptide substrate for the serine/threonine  
protein kinase family members, human YAK1, human YAK3a and hYAK3b, and yeast YAK1.  
More particularly, the peptide of the present invention is derived from the carboxyl terminal end of  
10 bovine myelin basic protein (MBP). It is a 15 amino acid peptide with Ser at position 164 of the  
bovine MBP sequence as the major site of phosphorylation by hYAK1, hYAK3a and hYAK3b, and  
yeast YAK1, hereinafter referred to as Ser164. The sequence of Ser164 is:

Leu-Gly-Gly-Arg-Asp-Ser-Arg-Ser-Gly-Ser-Pro-Met-Ala-Arg-Arg (SEQ ID NO:9)

155                      160                      164                      169

15 The invention also relates to a method of screening for inhibitors of human YAK1 and  
YAK3a and YAK3b, and yeast YAK1 protein serine/threonine kinases, and to inhibiting or  
activating the action of the above mentioned protein kinases and other YAK1 homologs by agents  
discovered using Ser164 or related peptide analogs as a method of screening for inhibitors and  
activators. In addition, this invention relates to using the Ser164 peptide to search for cellular  
20 substrates and/or cellular proteins that interact with the YAK family members of STPK's.

### Background of the Invention

A number of polypeptide growth factors and hormones mediate their cellular effects through a  
signal transduction pathway. Transduction of signals from the cell surface receptors for these ligands  
to intracellular effectors frequently involves phosphorylation or dephosphorylation of specific protein  
25 substrates by regulatory protein serine/threonine kinases (PSTK) and phosphatases. Serine/threonine  
phosphorylation is a major mediator of signal transduction in multicellular organisms. Receptor-  
bound, membrane-bound and intracellular PSTKs regulate cell proliferation, cell differentiation and  
signaling processes in many cell types.

Aberrant protein serine/threonine kinase activity has been implicated or is suspected in a  
30 number of pathologies such as rheumatoid arthritis, psoriasis, septic shock, bone loss, many cancers  
and other proliferative diseases. Accordingly, serine/threonine kinases and the signal transduction

pathways which they are part of are potential targets for drug design.

Several PSTKs are involved in regulation of cell cycling. One example is the cyclin-dependent kinases or CDKs (Peter and Herskowitz, Cell 1994: 79, 181-184), which are activated by binding to regulatory proteins called cyclins and control passage of the cell through specific cell cycle checkpoints. For example, CDK2 complexed with cyclin E allows cells to progress through the G1 to S phase transition. The complexes of CDKs and cyclins are subject to inhibition by low molecular weight proteins such as p16 (Serrano et al, Nature 1993: 366, 704), which binds to and inhibits CDK4. Deletions or mutations in p16 have been implicated in a variety of tumors (Kamb et al, Science 1994: 264, 436-440). Therefore, the proliferative state of cells and diseases associated with this state are dependent on the activity of CDKs and their associated regulatory molecules. In diseases such as cancer where inhibition of proliferation is desired, compounds that inhibit CDKs may be useful therapeutic agents. Conversely, activators of CDKs may be useful where enhancement of proliferation is needed, such as in the treatment of immunodeficiency.

The yeast YAK1 is a PSTK with sequence homology to CDKs. It was originally identified in *S. cerevisiae* as a mediator of cell cycle arrest caused by inactivation of the cAMP-dependent protein kinase PKA (Garrett et al, Mol Cell Biol. 1991: 11, 4045-4052). YAK1 kinase activity is low in cycling yeast but increases dramatically when the cells are arrested prior to the S-G2 transition. Increased expression of YAK1 causes growth arrest in yeast cells deficient in PKA. Therefore, YAK1 can act as a cell cycle suppressor in yeast.

Frequently, in diseases such as osteoporosis and osteoarthritis, patients have established lesions of bone or cartilage, respectively. Treatment of such lesions requires an agent that will stimulate new bone or cartilage formation to replace that lost to the disease; therefore, there is a need for drugs that increase the number of osteoblasts or chondrocytes, the cells responsible for bone or cartilage formation, respectively. Similarly, replacement of heart or skeletal muscle depleted by diseases such as myocardial infarction or HIV-associated cachexia requires drugs that stimulate proliferation of cardiac myocytes or skeletal myoblasts.

This indicates that these protein serine/threonine kinases have an established, proven history as therapeutic targets. Clearly there is a need for identification and characterization of further members of the protein serine/threonine kinase family which can play a role in preventing, ameliorating or correcting dysfunctions or diseases, including, but not limited to, bone loss including osteoporosis; inflammatory diseases such as Adult Respiratory Disease Syndrome (ARDS), Rheumatoid arthritis, Osteoarthritis, Inflammatory Bowel Disease (IBD), psoriasis, dermatitis, asthma, allergies; infections such as bacterial, fungal, protozoan and viral infections, particularly infections

caused by HIV-1 or HIV-2; HIV-associated cachexia and other immunodeficiency disorders; septic shock; pain; injury; cancers; anorexia; bulimia; Parkinson's disease; cardiovascular disease including restenosis, atherosclerosis, acute heart failure, myocardial infarction; hypotension; hypertension; urinary retention; angina pectoris; ulcers; benign prostatic hypertrophy; and psychotic and  
5 neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome.

Previously, a novel human homolog of yeast YAK1, termed hYAK1, which is expressed in osteoblasts, chondrocytes, cardiac and skeletal muscle, and at lower levels, in placenta and pancreas was described (CREASY et al., EP Publication Number 98301124.8, published February 16, 1998,  
10 the disclosure of which is incorporated herein by reference in its entirety). The sequence of hYAK1 shares homology with predicted PSTK's from *C. elegans*, *S. pombe* and *S. cerevisiae* and has motifs associated with known protein kinases. Inhibitors of hYAK1 are expected to stimulate proliferation of cells in which it is expressed.

Another novel human homolog of yeast YAK1, termed hYAK3b, which is expressed in testis,  
15 skeletal muscle, and in hematopoietic cells that can undergo erythroid differentiation (XIE, and CREASY, U.S. Patent Number 5,965,420, granted October 12, 1999, and U.S. Application Serial Number 09/359,257, filed July 22, 1999, the disclosures of which are incorporated herein by reference in their entireties). The sequence of hYAK3b shares 79% amino acid identity with hYAK1 in its kinase domain, and shares homology with predicted PSTK's from *C. elegans*, *S. pombe* and *S.*  
20 *cerevisiae*.

hYAK1 and hYAK3b are members of a novel subfamily of protein kinases with unique structural and enzymatic features. This family had been termed Dyrk (for dual specificity YAK-related kinases) by Dr. Joost and his co-workers. In their publications, hYAK1 and hYAK3b were termed Dyrk2 and Dyrk3, respectively (Becker and Joost, Prog. Nucl. Acid Res., 1999, 62, 1-17).

## 25 Summary of the Invention

In one aspect, the invention relates to a peptide substrate for hYAK1, hYAK3a and hYAK3b and yeast YAK1. This peptide, which has the sequence LGGRDSRSGSPMARR (SEQ ID NO:9), is derived from the carboxyl terminal end of bovine MBP (residues 155-169), and in it the Ser residue at position 164 (underlined) was identified as the major phosphoacceptor residue for  
30 hYAK1. Hereinafter this peptide is referred to as Ser164. This same residue was identified as the major phosphoacceptor site also for hYAK3a and 3b and yeast YAK1 PSTKs. One aspect of the invention relates to methods for using the Ser164 peptide as a substrate to screen for inhibitors of the kinase activity of hYAK1, hYAK3a, hYAK3b, yeast YAK1, and other members of this family of

protein kinases. Such uses include the treatment of bone loss including osteoporosis; inflammatory diseases such as Adult Respiratory Disease Syndrome (ARDS), Rheumatoid arthritis, Osteoarthritis, Inflammatory Bowel Disease (IBD), psoriasis, dermatitis, asthma, allergies; infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; HIV-associated cachexia and other immunodeficiency disorders; septic shock; pain; injury; cancers; anorexia; bulimia; Parkinson's disease; cardiovascular disease including restenosis, atherosclerosis, acute heart failure, myocardial infarction; hypotension; hypertension; urinary retention; angina pectoris; ulcers; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome, among others. In still another aspect, the invention relates to methods to identify agonists and antagonists using the materials provided by the invention, and treating conditions associated with hYAK1, hYAK3a and 3b, and yeast YAK1 imbalance with the identified compounds. Yet another aspect of the invention relates to diagnostic assays for detecting diseases associated with inappropriate hYAK1, hYAK3a and 3b, and yeast YAK1 activity or levels. Still another aspect of the invention is the use of the S164 peptide sequence to identify cellular proteins that are phosphorylated by or interact with hYAK1, hYAK3a and 3b, and yeast YAK1.

#### **Brief Description of the Drawings**

Figure 1 shows the determination of hYAK1-catalyzed phosphorylation sites on bovine MBP.

Figure 2 shows the activity of hYAK1, hYAK3b, and yeast YAK1 against bovine MBP and the S164 peptide.

Figure 3 shows the results of steady state two substrate analysis to determine the kinetic constants of the kinase reaction of purified recombinant hYAK1 using the Ser164 as substrate.

#### **Description of the Invention**

##### **Definitions**

The following definitions are provided to facilitate understanding of certain terms used frequently herein.

"hYAK1" refers, among others, generally to a polypeptide having the amino acid sequence set forth in SEQ ID NO:4 or an allelic variant thereof.

"hYAK1 activity or hYAK1 polypeptide activity" or "biological activity of the hYAK1 or hYAK1 polypeptide" refers to the metabolic or physiologic function of said hYAK1 including similar activities or improved activities or these activities with decreased undesirable side-effects. Also included are antigenic and immunogenic activities of said hYAK1.

5 "hYAK1 gene" refers to a polynucleotide having the nucleotide sequence set forth in SEQ ID NO:3 or allelic variants thereof and/or their complements.

"hYAK3a" refers, among others, generally to a polypeptide having the amino acid sequence set forth in SEQ ID NO:6 or an allelic variant thereof.

10 "hYAK3a activity or hYAK3a polypeptide activity" or "biological activity of the hYAK3a or hYAK3a polypeptide" refers to the metabolic or physiologic function of said hYAK3a including similar activities or improved activities or these activities with decreased undesirable side-effects. Also included are antigenic and immunogenic activities of said hYAK3a.

"hYAK3a gene" refers to a polynucleotide having the nucleotide sequence set forth in SEQ ID NO:5 or allelic variants thereof and/or their complements.

15 "hYAK3b" refers, among others, generally to a polypeptide having the amino acid sequence set forth in SEQ ID NO:8 or an allelic variant thereof.

"hYAK3b activity or hYAK3b polypeptide activity" or "biological activity of the hYAK3b or hYAK3b polypeptide" refers to the metabolic or physiologic function of said hYAK3b including similar activities or improved activities or these activities with decreased undesirable side-effects. 20 Also included are antigenic and immunogenic activities of said hYAK3b.

"hYAK3b gene" refers to a polynucleotide having the nucleotide sequence set forth in SEQ ID NO:7 or allelic variants thereof and/or their complements.

"Yeast YAK1" refers, among others, generally to a polypeptide having the amino acid sequence set forth in SEQ ID NO:2 or an allelic variant thereof.

25 "Yeast YAK1 activity or yYAK1 polypeptide activity" or "biological activity of the yYAK1 or yYAK1 polypeptide" refers to the metabolic or physiologic function of said yYAK1 including similar activities or improved activities or these activities with decreased undesirable side-effects. Also included are antigenic and immunogenic activities of said yYAK1.

30 "Yeast YAK1 gene" refers to a polynucleotide having the nucleotide sequence set forth in SEQ ID NO:1 or allelic variants thereof and/or their complements.

"Antibodies" as used herein includes polyclonal and monoclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, including the products of an Fab or other immunoglobulin expression library.

"Isolated" means altered "by the hand of man" from the natural state. If an "isolated" composition or substance occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living animal is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein.

"Polynucleotide" generally refers to any polyribonucleotide or polydeoxyribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. "Polynucleotides" include, without limitation single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, "polynucleotide" refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The term polynucleotide also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications has been made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short polynucleotides, often referred to as oligonucleotides.

"Polypeptide" refers to any peptide or protein comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres. "Polypeptide" refers to both short chains, commonly referred to as peptides, oligopeptides or oligomers, and to longer chains, generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene-encoded amino acids. "Polypeptides" include amino acid sequences modified either by natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well-described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide

may contain many types of modifications. Polypeptides may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. See, for instance, PROTEINS - STRUCTURE AND MOLECULAR PROPERTIES, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993 and Wold, F., Posttranslational Protein Modifications: Perspectives and Prospects, pgs. 1-12 in POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, New York, 1983; Seifter *et al.*, "Analysis for protein modifications and nonprotein cofactors", *Meth Enzymol* (1990) 182:626-646 and Rattan *et al.*, "Protein Synthesis: Posttranslational Modifications and Aging", *Ann NY Acad Sci* (1992) 663:48-62.

"Variant" as the term is used herein, is a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide respectively, but retains essential properties. A typical variant of a polynucleotide differs in nucleotide sequence from another, reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from another, reference polypeptide. Generally, differences are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, additions, or deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. A variant of a polynucleotide or polypeptide may be a naturally occurring such as an allelic variant, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis.



"Identity" reflects a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, determined by comparing the sequences. In general, identity refers to an exact nucleotide to nucleotide or amino acid to amino acid correspondence of the two polynucleotide or two polypeptide sequences, respectively, over the length of the sequences being compared. For sequences where there is not an exact correspondence, a "% identity" may be determined. In general, the two sequences to be compared are aligned to give a maximum correlation between the sequences. This may include inserting "gaps" in either one or both sequences, to enhance the degree of alignment. A % identity may be determined over the whole length of each of the sequences being compared (so-called global alignment), that is particularly suitable for sequences of the same or very similar length, or over shorter, defined lengths (so-called local alignment), that is more suitable for sequences of unequal length.

"Similarity" is a further, more sophisticated measure of the relationship between two polypeptide sequences. In general, "similarity" means a comparison between the amino acids of two polypeptide chains, on a residue by residue basis, taking into account not only exact correspondences between a between pairs of residues, one from each of the sequences being compared (as for identity) but also, where there is not an exact correspondence, whether, on an evolutionary basis, one residue is a likely substitute for the other. This likelihood has an associated 'score' from which the "% similarity" of the two sequences can then be determined.

Methods for comparing the identity and similarity of two or more sequences are well known in the art. Thus for instance, programs available in the Wisconsin Sequence Analysis Package, version 9.1 (Devereux J., *et al*, *Nucleic Acids Res*, 12, 387-395, 1984, available from Genetics Computer Group, Madison, Wisconsin, USA), for example the programs BESTFIT and GAP, may be used to determine the % identity between two polynucleotides and the % identity and the % similarity between two polypeptide sequences. BESTFIT uses the "local homology" algorithm of Smith and Waterman (*J. Mol. Biol.*, 147:195-197, 1981, *Advances in Applied Mathematics*, 2, 482-489, 1981) and finds the best single region of similarity between two sequences. BESTFIT is more suited to comparing two polynucleotide or two polypeptide sequences that are dissimilar in length, the program assuming that the shorter sequence represents a portion of the longer. In comparison, GAP aligns two sequences, finding a "maximum similarity", according to the algorithm of Needleman and Wunsch (*J. Mol. Biol.*, 48, 443-453, 1970). GAP is more suited to comparing sequences that are approximately the same length and an alignment is expected over the entire length. Preferably, the parameters "Gap Weight" and "Length Weight" used in each program are 50 and 3, for polynucleotide sequences and 12 and 4 for polypeptide sequences, respectively.

Preferably, % identities and similarities are determined when the two sequences being compared are optimally aligned.

Other programs for determining identity and/or similarity between sequences are also known in the art, for instance the BLAST family of programs (Altschul S.F., *et al.*, *J. Mol. Biol.*, 215, 403-410, 1990, Altschul S.F., *et al.*, *Nucleic Acids Res.*, 25:389-3402, 1997, available from the  
5 National Center for Biotechnology Information (NCBI), Bethesda, Maryland, USA and accessible through the home page of the NCBI at [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) and FASTA (Pearson W R, *Methods in Enzymology*, 183: 63-99 (1990); Pearson W R and Lipman D.J., *Proc Nat Acad Sci USA*, 85: 2444-2448 (1988) (available as part of the Wisconsin Sequence Analysis Package).

10 Preferably, the BLOSUM62 amino acid substitution matrix (Henikoff S. and Henikoff J.G., *Proc. Nat. Acad Sci. USA*, 89: 10915-10919 (1992)) is used in polypeptide sequence comparisons including where nucleotide sequences are first translated into amino acid sequences before comparison.

Preferably, the program BESTFIT is used to determine the % identity of a query  
15 polynucleotide or a polypeptide sequence with respect to a polynucleotide or a polypeptide sequence of the present invention, the query and the reference sequence being optimally aligned and the parameters of the program set at the default value, as hereinbefore described.

Alternatively, for instance, for the purposes of interpreting the scope of a claim including mention of a "% identity" to a reference polynucleotide, a polynucleotide sequence having, for example, at  
20 least 95% identity to a reference polynucleotide sequence is identical to the reference sequence except that the polynucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference sequence. Such point mutations are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, or insertion. These point mutations may occur at the 5' or 3' terminal positions of the reference polynucleotide  
25 sequence or anywhere between these terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence. In other words, to obtain a polynucleotide sequence having at least 95% identity to a reference polynucleotide sequence, up to 5% of the nucleotides of the in the reference sequence may be deleted, substituted or inserted, or any combination thereof, as hereinbefore described. The  
30 same applies *mutatis mutandis* for other % identities such as 96%, 97%, 98%, 99% and 100%.

For the purposes of interpreting the scope of a claim including mention of a "% identity" to a reference polypeptide, a polypeptide sequence having, for example, at least 95% identity to a reference polypeptide sequence is identical to the reference sequence except that the polypeptide

sequence may include up to five point mutations per each 100 amino acids of the reference sequence. Such point mutations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion. These point mutations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between these terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence. In other words, to obtain a sequence polypeptide sequence having at least 95% identity to a reference polypeptide sequence, up to 5% of the amino acids of the in the reference sequence may be deleted, substituted or inserted, or any combination thereof, as hereinbefore described. The same applies *mutatis mutandis* for other % identities such as 96%, 97%, 98%, 99%, and 100%.

A preferred meaning for "identity" for polynucleotides and polypeptides, as the case may be, are provided in (1) and (2) below.

(1) Polynucleotide embodiments further include an isolated polynucleotide comprising a polynucleotide sequence having at least a 95, 97 or 100% identity to the reference sequence of SEQ ID NO:1, (or SEQ ID NOS: 3, 5, or 7) wherein said polynucleotide sequence may be identical to the reference sequence of SEQ ID NO:1 or may include up to a certain integer number of nucleotide alterations as compared to the reference sequence, wherein said alterations are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, or insertion, and wherein said alterations may occur at the 5' or 3' terminal positions of the reference nucleotide sequence or anywhere between those terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence, and wherein said number of nucleotide alterations is determined by multiplying the total number of nucleotides in SEQ ID NO:1 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of nucleotides in SEQ ID NO:1, or:

$$n_n \leq x_n - (x_n \bullet y),$$

wherein  $n_n$  is the number of nucleotide alterations,  $x_n$  is the total number of nucleotides in SEQ ID NO:1,  $y$  is 0.95 for 95%, 0.97 for 97% or 1.00 for 100%, and  $\bullet$  is the symbol for the multiplication operator, and wherein any non-integer product of  $x_n$  and  $y$  is rounded down to the nearest integer prior to subtracting it from  $x_n$ . Alterations of a polynucleotide sequence encoding the polypeptide of SEQ ID NO:2 may create nonsense, missense or frameshift mutations in this coding sequence and thereby alter the polypeptide encoded by the polynucleotide following such alterations.

(2) Polypeptide embodiments further include an isolated polypeptide comprising a polypeptide having at least a 95, 97 or 100% identity to a polypeptide reference sequence of SEQ ID NO:2, (or SEQ ID NOS: 4, 6, or 8-18) wherein said polypeptide sequence may be identical to the reference sequence of SEQ ID NO:2 or may include up to a certain integer number of amino acid alterations as compared to the reference sequence, wherein said alterations are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion, and wherein said alterations may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between those terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence, and wherein said number of amino acid alterations is determined by multiplying the total number of amino acids in SEQ ID NO:2 by the integer defining the percent identity divided by 100 and then subtracting that product from said total number of amino acids in SEQ ID NO:2, or:

$$n_a \leq x_a - (x_a \bullet y),$$

wherein  $n_a$  is the number of amino acid alterations,  $x_a$  is the total number of amino acids in SEQ ID NO:2,  $y$  is 0.95 for 95%, 0.97 for 97% or 1.00 for 100%, and  $\bullet$  is the symbol for the multiplication operator, and wherein any non-integer product of  $x_a$  and  $y$  is rounded down to the nearest integer prior to subtracting it from  $x_a$ .

### Peptides of the Invention

In one aspect, the present invention relates to the Ser164 peptides. The Ser164 peptides include the peptide of the sequence in Figure 3; as well as peptide analogs which were made to study the structure activity relationship of Ser164 phosphorylation by these PSTK's. These analogs include but are not limited to peptides 1 to 8 below (modified residues are bolded):

Leu-Gly-Gly-Arg-Asp-Ser-Arg-Ser-Gly-Ser-Pro-Met-Ala-Arg-Arg (Ser164) (SEQ ID NO:9)

Leu-Gly-Gly-Arg-Asp-Ser-Arg-Ser-Gly-Ser-**Ala**-Met-Ala-Arg-Arg (Peptide 1) (SEQ ID NO:10)

Leu-Gly-Gly-Arg-Asp-Ser-Arg-Ser-Gly-**Thr**-Pro-Met-Ala-Arg-Arg (Peptide 2) (SEQ ID NO:11)

Leu-Gly-Gly-Arg-Asp-Ser-Arg-Ser-Gly-Ser-Pro-**Pro**-Ala-Arg-Arg (Peptide 3) (SEQ ID NO:12)

Leu-Gly-Gly-Arg-Asp-Ser-Arg-Ser-**Pro**-Ser-Pro-Met-Ala-Arg-Arg (Peptide 4) (SEQ ID NO:13)

Leu-Gly-Gly-Arg-Asp-Ser-Arg-Ser-**Pro**-Ser-**Ala**-Met-Ala-Arg-Arg (Peptide 5) (SEQ ID NO:14)

Leu-Gly-Gly-Arg-Asp-Ser-Arg-**Ala**-Gly-Ser-Pro-Met-Ala-Arg-Arg (Peptide 6) (SEQ ID NO:15)

Leu-Gly-Gly-Arg-Asp-Ala-Arg-Ser-Gly-Ser-Pro-Met-Ala-Arg-Arg (Peptide 7) (SEQ ID NO:16)

Leu-Gly-Gly-Arg-Asp-Ala-Arg-Ala-Gly-Ser-Pro-Met-Ala-Arg-Arg (Peptide 8) (SEQ ID NO:17)

Also included within the Ser164 peptides are truncated peptides analogs that do not have the first three residues (Leu-Gly-Gly).

5 **Detection of Ser164 as the site on MBP phosphorylated by the YAKs PSTKs:**

By screening several proteins as substrates, we determined that hYAK1, hYAK3a and 3b and yeast YAK1 preferred MBP as a substrate over casein, enolase, poly-Gly/Tyr, and several forms of histone. To identify the phosphorylation sites on bovine MBP catalyzed by each of the above mentioned protein kinases, bovine MBP was incubated for 30 minutes at 30 °C with 100 µM ATP, 10 mM MgCl<sub>2</sub>, and purified hYAK1, purified hYAK3b, or yeast YAK1 immunoprecipitated (using anti-HA Ab) from crude extracts of *Saccharomyces cerevisiae* transfected with HA-tagged yeast YAK1. Such an *in vitro* phosphorylated MBP was digested with endoproteinase Lys-C for 2 hours at 38 °C in 50 mM Tris pH 8.8. An aliquot of the unfractionated protein digest was analyzed by MALDI TOF for the presence of phosphorylated peptides. MALDI mass spectra (Annan, and Carr, 15 (1996) *Anal. Chem.* 68, 3413-3421) were recorded on a Micromass ToFSpec SE single-stage, gridless reflectron time-of-flight mass spectrometer equipped with a time-lag-focusing source. Samples were prepared by mixing the digest with an equal volume of matrix solution which contained two internal mass standards at a concentration of 200 fmol/µl. A 0.5 µl aliquot of this sample/matrix solution was applied to the MALDI target. The matrix solution was 10 µg/ml α-Cyano-4-hydroxycinnamic acid in 50:50 ethanol:acetonitrile. Samples were then irradiated by 337 nm photons from a pulsed Laser Science nitrogen laser operating at 5 Hz. Typically, 20-50 laser shots were summed into a single mass spectrum. Potential phosphopeptide precursor ions were selected for post-source-decay (PSD) analysis using a Bradbury-Nielsen ion gate. Phosphopeptides were isolated by reversed phase HPLC using a 1mm i.d. C18 column. Peptides are eluted using 20 acetonitrile:water:0.1% trifluoroacetic acid gradients at 50 µl/min. The column eluent was split post detector such that 5 µl/min was sent to a Micromass QTOF quadrupole time-of-flight mass spectrometer, and 45 µl/min went to a fraction collector taking one minute fractions. Phosphopeptides were sequenced by tandem MS (Verma, et al. (1997) *Science*, 278, 455-560.) using the QTOF mass spectrometer equipped with a nanoelectrospray source (Wilm, and Mann, 25 (1994) *Analytical Chem.* 66, 4390-4399 1994).

MALDI-TOF analysis of 5% of the unfractionated endoproteinase Lys-C digests from phosphorylated and nonphosphorylated MBP revealed 2 apparent phosphate-containing peptides at

m/z 1571.78 (MH<sup>+</sup>) and 1682.80 (MH<sup>+</sup>) (Fig. 1(labelled A)), corresponding to peptides of molecular mass 1570.77 and 1681.79, respectively. The 1570 Da species is consistent with the mass of a peptide containing residues 91-104 from MBP plus one mole of phosphate (calculated mass, 1570.78 Da), whereas the 1681 Da peptide was consistent with the mass of the C-terminal Lys-C fragment from MBP incorporating residues 155-169 plus one mole of phosphate (calculated mass, 1681.78). Observation of the characteristic loss of H<sub>3</sub>PO<sub>4</sub> (98 Da) from both peptides upon subsequent PSD analysis (Annan, and Carr, (1996) *Anal. Chem.* **68**, 3413-3421) confirmed that these peptides were, in fact, phosphorylated.

The Lys-C digests were fractionated by reverse phase HPLC using MS as an on-line readout of the fractions. Fractions containing the two putative phosphopeptides were analyzed by nanoES, and the appropriate precursor ions sequenced by collision-induced dissociation tandem mass spectrometry. The spectrum obtained for the 1570.9 peptide (Fig. 1(labelled B)) is consistent with two phosphorylated forms of the sequence NIVTPRTPPPSQGK (SEQ ID NO:18), which corresponds to residues 91-104 of bovine MBP. The major phosphorylation site was found to be Thr97. The 181.04 Da mass difference between the b<sub>6</sub> and b<sub>7</sub> ions indicates the presence of phosphothreonine at this position. A very small amount of phosphorylation was also found at Thr94. The evidence for phosphorylation at this second site was a series of b<sub>n</sub>Δ ions starting at b<sub>4</sub> and ending at b<sub>6</sub>. (The evidence for phosphorylation at either site would converge at the b<sub>7</sub> ion.)

The tandem mass spectrum of the 1681 peptide (Fig. 1(labelled C)) is consistent with a monophosphorylated form of LGGRDSRSGSPMARR (SEQ ID NO:9), corresponding to the C-terminal residues 155-169 of bovine MBP. Although the tandem MS data clearly indicate that Ser160 is not phosphorylated, it was unable to distinguish between Ser162 or Ser164 as the sole site of phosphorylation in this peptide. Assignment of the phosphorylation site as Ser164 is based on the presence of an S/TP sequence at this site, which is analogous to the motif found at Thr94 and Thr97.

Phosphorylation of bovine MBP on Thr97 by the MAP kinases has been previously reported (Erikson, et al., (1990) *J. Biol. Chem.* **265**, 19728-19735; Clark-Lewis, et al., (1991) *J. Biol. Chem.* **266**, 11580-11584.). However, we are not aware of any kinase that phosphorylates bovine MBP on Ser164. Therefore, the phosphorylation pattern of yeast YAK1 on MBP appears to be unique.

#### **Kinase Assays Using Ser164 as the phosphoacceptor**

The source of Ser164 peptide: It was custom-ordered from California Peptide Research Inc. (Napa, CA), and its purity was determined by HPLC. The peptide contained 15 amino acids, and its

calculated molecular mass was 1601.82 daltons. Solid sample was dissolved at 5 mM in ice-cold kinase assay buffer (see later), aliquoted, and stored at  $-20^{\circ}\text{C}$  until use.

Figure 1 shows a determination of YAK1-induced phosphorylation sites on myelin basic protein. **A**, Partial MALDI mass spectrum of an unfractionated Lys-C digest of *in vitro* phosphorylated MBP. Peptides modified by a single mole of phosphate show an increase in mass of 80 Da. **B**, ES-CID tandem mass spectrum of the  $\text{M}^{2+}$  ion ( $m/z$  786.4) from the 1570 Da phosphorylated peptide shown above. Ions marked  $b_n\Delta$  have the structure  $b_n\text{-H}_3\text{PO}_4$ . **C**, ES-CID tandem mass spectrum of the  $\text{M}^{3+}$  ion ( $m/z$  561.6) from the 1681 Da phosphorylated peptide shown above. Ions marked  $b_n^{\Psi}\Delta$  have the structure  $b_n\text{-H}_3\text{PO}_4\text{-NH}_3$ .

Figure 2 shows the activity of hYAK1, hYAK3, and yeast YAK1 against MBP and the Ser164 peptide. 5 ng purified hYAK1 and 100 ng purified hYAK3 were used per assay. Anti-HA mAb immune complex kinase assay was performed on 100  $\mu\text{g}$  protein from crude extracts of yeast cells expressing either FL or  $\Delta\text{N}$  yeast YAK1. Concentration of ATP was 100  $\mu\text{M}$ , Ser164 was used at 0.5 mM, and MBP was at 10  $\mu\text{g}$ /reaction (18.5  $\mu\text{M}$ ).

Figure 3 shows double reciprocal plots of enzyme velocity vs. ATP or Ser164 concentration. Steady state two substrate analysis of purified hYAK1 using S164 as phosphoacceptor. 1 ng hYAK1 was incubated for 30 minutes at  $37^{\circ}\text{C}$  in hYAK1 reaction mix that contains various concentrations of ATP and Ser164. Each ATP concentration was used with 1, 0.75, 0.5, 0.33, 0.25, 0.15, and 0.05 mM Ser164. Each Ser164 concentration was used with 333, 166, 100, 66.7, 33.3, 16.7, and 6.6  $\mu\text{M}$  ATP. Reciprocal values of enzyme velocity were calculated and plotted against reciprocal values of the concentration of each substrate. Data were analyzed by GraphFit, and calculated kinetic constants are shown.

#### The source of enzyme:

1) hYAK1: DET1/DET2-tagged full length hYAK1 was expressed in *Drosophila* sf9 cells and purified to >95% purity using Ni column chromatography. The purified protein migrated on SDS gels as a single band with an apparent molecular mass of 62 kDa. Samples were stored at  $-80^{\circ}\text{C}$  until use.

2) hYAK3b: Glutathione-S-Transferase (GST)/Factor Xa-tagged hYAK3b was expressed in baculovirus cells and purified to about 50% purity using Glutathione Sepharose 4B column chromatography, followed by Ni-NTA column chromatography. Samples were stored at  $-80^{\circ}\text{C}$  until use.

3) Yeast YAK1: Full length and an amino-terminally truncated (amino acids 148-807, termed  $\Delta$ N) hemagglutinin (HA)-tagged yeast YAK1 was each expressed in a strain of *S. cerevisiae* lacking the endogenous *YAK1* gene and all three *PKA* genes. Cultures for experiments were grown in liquid Sc-His to an OD<sub>600</sub> of at least 1.0, washed with Sc-His g/r, resuspended in Sc-His g/r to twice the original volume and grown for 16-24 hours at RT. Cells were washed once with H<sub>2</sub>O and the pellets stored at -80 °C until use. To prepare lysates, cell pellets were thawed and resuspended at 1 ml/100 ml of original culture in lysis buffer (LB) containing 50 mM Tris pH 7.5, 150 mM NaCl, 10 µg/ml each aprotinin, leupeptin and TLCK, 0.1 mM PMSF, 50 mM NaF, 1 mM NaVanadate, 10 mM β-glycerophosphate. Following the addition of 0.5 ml sterile acid-washed glass beads, cells were disrupted via ten, 30 second intervals of vortexing. NP40 was added to a 2% final concentration followed by rocking at 4 °C for 30-50 minutes. Lysates were clarified by high-speed centrifugation, and the supernatants were stored at -80 °C until use. Each form of yeast YAK1 was immunoprecipitated from the detergent extracts using anti-HA mAb.

Immune Complex Protein Kinase Assay for Yeast YAK1: Yeast cellular extracts were immunoprecipitated by rocking overnight at 4°C with 4 µg of the anti-HA tag antibody and 100 µl of 20% suspension of protein A agarose (GIBCO-BRL) in LB that contained 1% NP-40. Samples were then washed twice with LB and once with basic kinase assay buffer (25 mM Hepes, pH 7.5; 1 mM DTT; 10 mM β-glycerol phosphate; 0.2 mM NaV). Washed immune complexes were suspended in 20 µl of basic kinase assay buffer that contained 0.1 mM ATP, 3 µCi of [γ-<sup>32</sup>P]ATP, 10 mM MgCl<sub>2</sub>, plus either bovine MBP or the Ser164 peptide. After incubation for 15 minutes at 30°C, the reactions were stopped by adding 20 µl of 0.15 M phosphoric acid. Phosphorylated substrates were isolated by spotting 20 µl of each sample on phosphocellulose (p81) filters. Filters were washed 3 times with 75 mM phosphoric acid followed by 3 times with H<sub>2</sub>O, and counted for <sup>32</sup>P incorporation using a β-scintillation counter.

Kinase assay of purified hYAK1 and hYAK3: The assay was performed in 96 well Minisorp plates (Costar, Catalog No 3356). Reaction (in 30 µl volume) mixtures contained in final concentrations, 25 mM Hepes buffer, pH 7.5; 0.2 mM sodium vandate; 10 mM MgCl<sub>2</sub>; 1 mM DTT; 10 mM β-glycerol phosphate; 0.1% BSA; 0.1 mM ATP, 2.5 µCi of [γ-<sup>32</sup>P]ATP; purified hYAK1 (1-5 ng/assay), or purified hYAK3b (50-100 ng/assay); and either bovine MBP or the Ser164 peptide used at the concentrations indicated below and in the descriptions of the figures. Reactions were incubated for 20 minutes at 37°C, and were stopped by adding 10 µl of 0.3 M phosphoric acid.



Phosphorylated substrates were isolated by spotting 20 µl of the reaction on p81 filters, and processed as detailed earlier.

This same assay can be performed on a FlashPlate format in which the plate is coated with MBP or with the S164 peptide by incubation overnight at 4 °C in 100 µl of either substrate dissolved in Sodium Carbonate buffer, pH 8.8. When coating with MBP, a solution of 100 µg/ml MBP was used to coat wells with 100 µl (10 µg) MBP per well. When coating with Ser164, a solution of 0.4 mg/ml (0.25 mM) was used to coat wells with 100 µl (40 µg) Ser164 per well. An example of a FlashPlate assay protocol and typical results are given below:

#### **FlashPlate (FP) Protocol**

1. Coat Maxisorp plates with MBP or Ser164 as above.
2. Wash plates once with kinase assay buffer (KB): 25 mM Hepes, pH 7.5; 0.2 mM NaV; 10 mM β-glycerol phosphate; 1 mM Na pyrophosphate
3. Add enzyme (Ni-hYAK1, diluted in KB), DMSO or inhibitors (in KB) and keep on ice for 30 minutes
4. Add KB containing Mg/ATP to a [final] of 0.1 mM [ $\gamma$ - $^{33}\text{P}$ ]ATP and 10 mM  $\text{MgCl}_2$
5. Incubate with shaking, 1-2 hrs, at room temperature
6. Aspirate and wash 6 times with 0.5 ml KB
7. Read  $^{33}\text{P}$  incorporation in FP reader
8. Blank = No enzyme added
9. Reaction volume: 25, 50 or 100 µl
10. 0.5 or 1.0 µCi  $^{33}\text{P}$ /0.1 mM ATP
11. MBP-FP better than basic FP (in house coating)
12. 37 °C incubation was not better (several time points)
13. Other incubation times at room temperature were not better

---

**Results:** Each kinase phosphorylated the Ser164 peptide to much higher specific activity than MBP (Figure 2). Steady state kinetic constants of the hYAK1 reaction were generated by

- 17 -

varying both substrates simultaneously and fitting enzyme velocity as a function of each substrate concentration. Double reciprocal plots ( $1/V$  vs.  $1/[\text{Substrate}]$ ) with S164 peptide as the phosphate acceptor are shown in Figure 3. GraphFit analysis of the results generated the following steady state kinetic constants:

- 5                     $K_m[\text{ATP}] = 42 \pm 7 \mu\text{M}$  .  
                        $K_m[\text{S164}] = 160 \pm 14 \mu\text{M}$ .  
                        $V_{max} = 51 \pm 6 \mu\text{mol/mg}$ .  
                        $k_{cat} = 160 \pm 19 \text{ min}^{-1}$ .

Typical results of FlashPlate are shown below

10    FlashPlate Typical results

Signal to noise ratio:            >7 fold

Blanks:                            30-80 cpm

[Ni-hYAK1]:                    As low as 20 ng/reaction (5 nM) for 100  $\mu\text{l}$  reactions  
    As low as 8 ng/reaction (5 nM) for 25  $\mu\text{l}$  reactions

15     $^{33}\text{P}$ :                            As low as 0.5  $\mu\text{Ci}$

Kinase inhibitors:            Potency comparable to tube assay:

SKF-108752 IC<sub>50</sub>:            0.19  $\mu\text{M}$  (0.1  $\mu\text{g}$  hYAK1)

0.13  $\mu\text{M}$ ; 0.16  $\mu\text{M}$  (0.3  $\mu\text{g}$  hYAK1)

K252a IC<sub>50</sub>:                    0.552  $\mu\text{M}$ ; 0.427  $\mu\text{M}$  (0.3  $\mu\text{g}$  hYAK1)

20    Specific Activity:            At 20 ng enzyme, MBP gave  $58 \pm 3$  ( $n = 6$ ), and Ser164 gave  $484 \pm 63$  ( $n = 6$ )  
    nmol/mg protein

DMSO:                            No effect up to 3%

Variability:                    <10% (between wells and from plate to plate)

Screening Assays

- 25                    The hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptides of the present invention may be employed in a screening process for compounds, which activate (agonists) or inhibit activation of (antagonists, or otherwise called inhibitors) the hYAK1, hYAK3a and 3b, and yeast YAK1

polypeptides of the present invention. Thus, polypeptides of the invention may also be used to assess or identify agonists or antagonists from, for example, cells, cell-free preparations, chemical libraries, and natural product mixtures. These agonists or antagonists may be natural substrates, ligands, receptors, etc., as the case may be, of the polypeptides of the present invention; or may be structural or functional mimetics of the polypeptide of the present invention. See Coligan *et al.*, *Current Protocols in Immunology* 1(2): Chapter 5 (1991).

hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptides are ubiquitous in the mammalian host and are responsible for many biological functions, including many pathologies. Accordingly, it is desirable to find compounds and drugs which stimulate hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptide on the one hand and which can inhibit the function of hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptide on the other hand. In general, agonists are employed for therapeutic and prophylactic purposes for such conditions as bone loss including osteoporosis; inflammatory diseases such as Adult Respiratory Disease Syndrome (ARDS), Rheumatoid arthritis, Osteoarthritis, Inflammatory Bowel Disease (IBD), psoriasis, dermatitis, asthma, allergies; infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; HIV-associated cachexia and other immunodeficiency disorders; septic shock; pain; injury; cancers; anorexia; bulimia; Parkinson's disease; cardiovascular disease including restenosis, atherosclerosis, acute heart failure, myocardial infarction; hypotension; hypertension; urinary retention; angina pectoris; ulcers; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome. Antagonists may be employed for a variety of therapeutic and prophylactic purposes for such conditions as bone loss including osteoporosis; inflammatory diseases such as Adult Respiratory Disease Syndrome (ARDS), Rheumatoid arthritis, Osteoarthritis, Inflammatory Bowel Disease (IBD), psoriasis, dermatitis, asthma, allergies; infections such as bacterial, fungal, protozoan and viral infections, particularly infections caused by HIV-1 or HIV-2; HIV-associated cachexia and other immunodeficiency disorders; septic shock; pain; injury; cancers; anorexia; bulimia; Parkinson's disease; cardiovascular disease including restenosis, atherosclerosis, acute heart failure, myocardial infarction; hypotension; hypertension; urinary retention; angina pectoris; ulcers; benign prostatic hypertrophy; and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome.

In general, such screening procedures may involve using appropriate cells which express the hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptides or respond to hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptide of the present invention. Such cells include cells from mammals, yeast,

*Drosophila* or *E.coli*. Cells which express the hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptides (or cell membrane containing the expressed polypeptides) or respond to hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptides are then contacted with a test compound to observe binding, or stimulation or inhibition of a functional response. The ability of the cells, which were contacted with the candidate compounds, is compared with the same cells which were not contacted for hYAK1, hYAK3a and 3b, and yeast YAK1 activity.

The knowledge that the hYAK1, hYAK3a and 3b, and yeast YAK1 encodes a protein kinase suggests that recombinant forms can be used to establish a protein kinase activity. Typically this would involve the direct incubation of hYAK1, hYAK3a and 3b, and yeast YAK1 with a protein or peptide substrate in the presence of  $\gamma$ -<sup>32</sup>P-ATP, followed by the measurement of radioactivity incorporated into the substrate by separation and counting. Separation methods include immunoprecipitation, conjugation of substrate to a bead allowing separation by centrifugation or determination of incorporation by scintillation proximity assay, SDS-PAGE followed by autoradiography or biosensor analysis. While the specific substrates are not yet known, candidates include hYAK1, hYAK3a and 3b, and yeast YAK1 themselves (autophosphorylation), myelin basic protein, casein, histone and HSP27. Other substances might be discovered by incubating hYAK1, hYAK3a and 3b, and yeast YAK1 with random peptides conjugated to solid supports or displayed on the surface of phage or by incubation of hYAK1, hYAK3a and 3b, and yeast YAK1 with mammalian cell lysates and  $\gamma$ -<sup>32</sup>P-ATP, followed by separation of the labelled target proteins, and sequencing. The protein kinase activity of hYAK1, hYAK3a and 3b, and yeast YAK1 may require incubation with a specific upstream effector. This may be achieved by preincubating hYAK1, hYAK3a and 3b, and yeast YAK1 with lysates from a variety of stimulated eukaryotic cells and ATP. These assays permit the discovery and modification of compounds which inhibit hYAK1, hYAK3a and 3b, and yeast YAK1 kinase activity in vitro and would be expected to have effects on proliferation of osteoblasts, chondrocytes, cardiac myocytes or skeletal myoblasts. Any inhibitors so identified would be expected to have up-regulatory effects on proliferation and be useful as a therapeutic for the treatment and prevention of diseases such as osteoporosis, osteoarthritis, cardiomyopathy and cachexia.

This invention contemplates the treatment and/or amelioration of such diseases by administering a hYAK1, hYAK3a and 3b, or yeast YAK1 inhibiting amount of a compound. Without wishing to be bound by any particular theory of the functioning of the hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptides of this invention, it is believed that among the useful inhibitors of hYAK1, hYAK3a and 3b, and yeast YAK1 function are those compounds which inhibit the kinase activity of the hYAK1, hYAK3a and 3b, and yeast YAK1. Other sites of inhibition are, of course, possible

owing to its position in a signal transduction cascade. Therefore, inhibiting the interaction of hYAK1, hYAK3a and 3b, and yeast YAK1 with one or more of its upstream or downstream modulators/substrates is also contemplated by this invention. Inhibitors of protein-protein interactions between hYAK1, hYAK3a and 3b, and yeast YAK1 and other factors could lead to the development of pharmaceutical agents for the modulation of hYAK1, hYAK3a and 3b, and yeast YAK1 activity.

The assays may simply test binding of a candidate compound wherein adherence to the cells bearing the hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptide is detected by means of a label directly or indirectly associated with the candidate compound or in an assay involving competition with a labeled competitor. Further, these assays may test whether the candidate compound results in a signal generated by activation of the hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptide, using detection systems appropriate to the cells bearing the hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptide. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed. Standard methods for conducting such screening assays are well understood in the art.

Examples of potential hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptide antagonists include antibodies or, in some cases, oligonucleotides or proteins which are closely related to the ligands, substrates, receptors, etc., as the case may be, of the hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptide, e.g., a fragment of the ligands, substrates, receptors, or small molecules which bind to the polypeptide of the present invention but do not elicit a response, so that the activity of the polypeptide is prevented.

#### Prophylactic and Therapeutic Methods

This invention provides methods of treating an abnormal condition related to both an excess of and insufficient amounts of hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptide activity.

If the activity of hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptides is in excess, several approaches are available. One approach comprises administering to a subject an inhibitor compound (antagonist) as hereinabove described along with a pharmaceutically acceptable carrier in an amount effective to inhibit activation by blocking binding of ligands to the hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptides, or by inhibiting a second signal, and thereby alleviating the abnormal condition.

In another approach, soluble forms of hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptides still capable of binding the ligand in competition with endogenous hYAK1, hYAK3a

and 3b, and yeast YAK1 polypeptides may be administered. Typical embodiments of such competitors comprise fragments of the hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptides.

In still another approach, expression of the gene encoding endogenous hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptide can be inhibited using expression blocking techniques. Known  
5 such techniques involve the use of antisense sequences, either internally generated or separately administered. See, for example, O'Connor, *J Neurochem* (1991) 56:560 in Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988). Alternatively, oligonucleotides which form triple helices with the gene can be supplied. See, for example, Lee *et al.*, *Nucleic Acids Res* (1979) 6:3073; Cooney *et al.*, *Science* (1988) 241:456; Dervan *et al.*, *Science*  
10 (1991) 251:1360. These oligomers can be administered *per se* or the relevant oligomers can be expressed *in vivo*.

For treating abnormal conditions related to an under-expression of hYAK1, hYAK3a and 3b, and yeast YAK1 and its activity, several approaches are also available. One approach comprises administering to a subject a therapeutically effective amount of a compound which activates hYAK1,  
15 hYAK3a and 3b, and yeast YAK1 polypeptide, i.e., an agonist as described above, in combination with a pharmaceutically acceptable carrier, to thereby alleviate the abnormal condition. Alternatively, gene therapy may be employed to effect the endogenous production of hYAK1, hYAK3a and 3b, and yeast YAK1 by the relevant cells in the subject. For example, a polynucleotide of the invention may be engineered for expression in a replication defective retroviral vector, as discussed above. The  
20 retroviral expression construct may then be isolated and introduced into a packaging cell transduced with a retroviral plasmid vector containing RNA encoding a polypeptide of the present invention such that the packaging cell now produces infectious viral particles containing the gene of interest. These producer cells may be administered to a subject for engineering cells *in vivo* and expression of the polypeptide *in vivo*. For overview of gene therapy, see Chapter 20, *Gene Therapy and other*  
25 *Molecular Genetic-based Therapeutic Approaches*, (and references cited therein) in Human Molecular Genetics, T Strachan and A P Read, BIOS Scientific Publishers Ltd (1996).

#### Formulation and Administration

Peptides, such as the soluble form of hYAK1, hYAK3a and 3b, and yeast YAK1 polypeptides, and agonists and antagonist peptides or small molecules, may be formulated in  
30 combination with a suitable pharmaceutical carrier. Such formulations comprise a therapeutically effective amount of the polypeptide or compound, and a pharmaceutically acceptable carrier or excipient. Such carriers include but are not limited to, saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The formulation should suit the mode of administration,

and is well within the skill of the art. The invention further relates to pharmaceutical packs and kits comprising one or more containers filled with one or more of the ingredients of the aforementioned compositions of the invention.

Polypeptides and other compounds of the present invention may be employed alone or in conjunction with other compounds, such as therapeutic compounds.

Preferred forms of systemic administration of the pharmaceutical compositions include injection, typically by intravenous injection. Other injection routes, such as subcutaneous, intramuscular, or intraperitoneal, can be used. Alternative means for systemic administration include transmucosal and transdermal administration using penetrants such as bile salts or fusidic acids or other detergents. In addition, if properly formulated in enteric or encapsulated formulations, oral administration may also be possible. Administration of these compounds may also be topical and/or localized, in the form of salves, pastes, gels and the like.

The dosage range required depends on the choice of peptide, the route of administration, the nature of the formulation, the nature of the subject's condition, and the judgment of the attending practitioner. Suitable dosages, however, are in the range of 0.1-100  $\mu\text{g/kg}$  of subject. Wide variations in the needed dosage, however, are to be expected in view of the variety of compounds available and the differing efficiencies of various routes of administration. For example, oral administration would be expected to require higher dosages than administration by intravenous injection. Variations in these dosage levels can be adjusted using standard empirical routines for optimization, as is well understood in the art.

Polypeptides used in treatment can also be generated endogenously in the subject, in treatment modalities often referred to as "gene therapy" as described above. Thus, for example, cells from a subject may be engineered with a polynucleotide, such as a DNA or RNA, to encode a polypeptide *ex vivo*, and for example, by the use of a retroviral plasmid vector. The cells are then introduced into the subject.

### Examples

The examples below are carried out using standard techniques, which are well known and routine to those of skill in the art, except where otherwise described in detail. The examples illustrate, but do not limit the invention.

#### Example 1

A partial clone (ATG-355, HGS EST # 454640) was initially identified through random searches of the Human Genome Sciences database. This partial clone (~ 1kb) showed significant

- 23 -

homology to YAK1 from *S. cerevisiae*. To get the full length cDNA: A total of 1M plaques were screened from a Human Osteoblast cDNA library (Stratagene, LaJolla CA) using the insert of the above partial clone as a probe. Library screening procedure is described by (Elgin, et al. Strategies 4: 8-9, 1991). The probes were  $\alpha$ -<sup>32</sup>P labeled, using Random Primed Labeling Kit (Boehringer Mannheim, Germany, Cat. # 1585584 ) and purified by running over Sephadex G-50 columns (Pharmacia Biotech. Cat. # 17-0855-02) The hybridization and washing conditions were according to ( J. Sambrook, E.F. Fritsch and T. Maniatis (1989) A Laboratory Manual Second. Ed. Vol. 1 pp. 2.69-2.81 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York). Five clones were isolated by plaque purification and a fragments containing the inserts were excised and sequenced. The longest insert was a 2.7 kb fragment containing the 3' untranslated region and part of the coding sequence of hYAK1. A second probe prepared by EcoRI-BamHI digestion of this insert followed by labelling with <sup>32</sup>P was used to screen a commercially available human heart cDNA library (Stratagene #936207). An additional five clones were plaque purified, excised into phagemids and sequenced. Fasta analysis show this peptide to have high homology to a putative serine/threonine kinase of unknown function from *C. elegans* (F49E11.1).

To confirm that the cDNA was full length, Human Leukocyte "Marathon Ready" cDNA (Clontech, Palo Alto, CA) was used as a template to amplify a fragment corresponding to the 5' region of hYAK1 using a 5' anchor primer-1 (Clontech) and a reverse gene specific primer. This fragment was T/A cloned into pCR2.1 (Invitrogen), and multiple isolates were sequenced. An in-frame stop codon was identified upstream of the predicted initiation codon confirming that the full-length cDNA had been obtained.

Northern analysis was carried out to determine the distribution of hYAK1 mRNA in human tissues. A fragment containing the 3' untranslated region of hYAK1 was isolated from SEQ ID NO:1 using standard techniques. The isolated fragment was radiolabelled with  $\alpha$ -<sup>32</sup>P-dATP using a randomly primed labelling kit. Membranes containing mRNA from multiple human tissues (Clontech #7760-1) were hybridized with the probe and washed under high stringency conditions as directed. Hybridized mRNA was visualized by exposing the membranes for 4 days to X-ray film. Three major transcripts of 2.6, 7 and 10 kb were present and were expressed most prominently in heart and skeletal muscle, but were present to a lesser degree in pancreas, placenta and brain. All three transcripts appeared absent from kidney.



## SEQUENCE INFORMATION

## SEQ ID NO:1

ATGAACTCAT CCAATAATAA CGACTCGTCC AGCTCCAATA GCAACATGAA TAACTCCTTG  
AGCCCGACCC TTGTGACCCA CAGTGATGCT AGTATGGGCT CGGGTAGAGC AAGCCCAGAC  
5 AATAGCCATA TGGGGAGAGG TATATGGAAT CCATCGTACG TAAATCAAGG CTCTCAAAGG  
TCTCCACAGC AGCAGCATCA GAATCATCAC CAGCAGCAGC AGCAGCAGCA ACAACAACAA  
CAACAGAATT CTCAATTTTG CTTTGTTAAC CCTTGGAATG AGGAAAAAGT AACTAATTCT  
CAACAAAACC TGGTGTATCC CCCTCAATAC GATGACTTAA ACAGTAACGA AAGTCTAGAT  
GCGTACAGAC GACGTAAATC TAGTCTCGTT GTACCTCCAG CCAGGGCACC TGCCCCAAAT  
10 CCTTTCAGT ACGATAGTTA TCCCGCTTAC ACCAGCTCTA ATACGAGTTT GGCAGGAAAT  
AGCAGTGGCC AGTATCCTTC TGGCTATCAA CAACAACAAC AGCAAGTATA CCAGCAGGGC  
GCTATCCATC CTTCCCAATT TGGATCCAGA TTTGTTCCCT CCCTTTATGA TCGTCAAGAT  
TTCCAAAGAA GGCAGAGTCT GGCTGCAACT AATTATTCGT CCAATTTTTC TTCCCTAAAT  
TCAAATACTA ATCAAGGAAC AAATCCATA CCAGTCATGT CACCTTACAG GAGGCTTAGC  
15 GCATATCCTC CTAGTACAAG TCCCCACTA CAACCCCTT TCAAGCAGTT ACGAAGAGAT  
GAAGTACAAG GTCAGAAATT GTCCATTCCT CAGATGCAAC TTTGCAACTC TAAAAACGAT  
CTTCAACCTG TCTTAAATGC TACTCCAAAG TTTAGACGTG CGTCGCTCAA TTCCAAGACG  
ATATCTCCCC TAGTCAGTGT CACGAAAAGC CTGATTACCA CATATTCTTT ATGTTCTCCA  
GAATTTACCT ACCAAACGTC CAAAAACCCC AAGAGAGTAC TTACGAAGCC CAGTGAAGGG  
20 AAATGTAACA ACGGATTCGA CAACATAAAC AGTGA CTATA TTCTTTATGT AAATGACGTT  
TTGGGTGTAG AGCAGAACAG AAAGTACCTT GTGCTAGACA TTTTGGGTCA AGGTACATTT  
GGTCAAGTGG TCAATGTCA AAATTGTCTG ACGAAAGAGA TATTGGCTGT AAAAGTGGTT  
AAATCGAGGA CAGAATATTT GACTCAAAGT ATAACGGAGG CTAAAATTTT AGAGCTACTG  
AATCAAAAGA TAGACCCTAC TAATAAACAT CATTTTTTAA GGATGTATGA CTCCTTTGTC  
25 CACAAGAACC ACTTATGTTT AGTGTGTTGAA TTAATAAGCA ATAATTTGTA CGAATTATTA  
AAACAAAACA AATTTTCATGG GCTCTCTATA CAACTTATCA GAACATTCAC CACTCAAATA  
CTGGATTCGT TATGCGTTTT GAAGGAAAGC AAATGATTC ATTGCGATCT GAAGCCAGAG  
AATATCTTGC TCTGCGCACC TGATAAGCCG GAATTGAAAA TTATTGATTT TGGCTCATCT  
TGCGAAGAGG CAAGAACCGT TTATACATAC ATCCAGTCCA GGTTTTATCG TGCTCCTGAA  
30 ATTATACTGG GTATACCGTA TTCAACCACT ATTGACATGT GGTCGTTAGG TTGTATTGTT  
GCTGAATTAT TTTTGGGTAT ACCGATCTTC CCAGGCGCTT CTGAATATAA CCAATTAACA  
AGAATAATAG ACACGCTTGG ATATCCTCCA TCGTGGATGA TAGATATGGG TAAAAACTCT  
GGAAAATTTA TGAAGAAATT GGCACCAGAA GAAAGTTCTT CTTCTACACA AAAGCATCGT  
ATGAAAAC TAAGAAGAGTT TTGCAGAGAA TACAATATAG TGGAAAAGCC CAGTAAACAA  
35 TATTTTAAGT GGAGAAAGTT ACCAGATATT ATTAGAACT ACAGGTATCC TAAAAGCATA  
CAGAACTCCC AAGAACTTAT CGACCAAGAA ATGCAGAATA GGGAGTGTTT GATCCACTTT  
TTAGGCGGTG TGCTAAATTT GAACCCGTTA GAAAGATGGA CACCACAACA AGCTATGCTA  
CACCCCTTCA TAACAAAGCA GGAGTTTACA GGTGAGTGGT TTCCTCCAGG ATCGTCTTTA  
CCGGGTCCTT CAGAAAAACA TGACGATGCA AAAGGCCAGC AAAGTGAATA TGGAAGTGCG  
40 AACGACTCTA GTAACAATGC AGGCCACAAC TATGTCTATA ATCCTAGCTC TGCCACTGGT

- 25 -

GGTGCTGATA GCGTCGACAT TGGTGCTATC AGTAAAAGGA AGGAGAATAC ATCTGGCGAC  
ATCTCCAATA ATTTTGCTGT TACTCATTCT GTTCAAGAAG GGCCAACAAG CGCGTTCAAT  
AAACTTCACA TTGTGAAGA ATAA

## SEQ ID NO:2

5 MNSSNNNDSS SSNSNMNNSL SPTLVTHSDA SMGSGRASPD NSHMGRGIWN PSYVNQGSQR  
SPQQQHQNHH QQQQQQQQQQ QONSQFCFVN PWNEEKVTNS QQNLVYPPQY DDLNSNESLD  
AYRRRKSSLV VPPARAPAPN PFQYDSYPAY TSSNTSLAGN SSGQYPSGYQ QQQQVYQQG  
AIHPSQFGSR FVPSLYDRQD FQRRQSLAAT NYSSNFSSLN SNTNQGTNSI PVMSPYRRLS  
AYPPSTSPPL QPPFKQLRRD EVQGQKLSIP QMQLCNSKND LQPVLNATPK FRRASLNSKT  
10 ISPLVSVTKS LITTYSLCSP EFTYQTSKNP KRVLTKPSEG KCNNGFDNIN SDYILYVNDV  
LGVEQNRKYL VLDILGQGTG GQVVKCQNL TKEILAVKV KSRTEYLTQS ITEAKILELL  
NQKIDPTNKH HFLRMYDSFV HKNHLCCLVFE LLSNNLYELL KQNKFHGLSI QLIRFTTTQI  
LDSLCLVKES KLIHCDLKPE NILLCAPDKP ELKIIDFGSS CEEARTVYTY IQSRFYRAPE  
IILGIPYSTS IDMWSLGCIV AELFLGIPIF PGASEYNQLT RIIDTLGYPP SWMIDMGKNS  
15 GKFMKKLAPE ESSSTQKHR MKTIEEFCRE YNIVEKPSKQ YFKWRKLPDI IRNYRYPKSI  
QNSQELIDQE MQNRECLIHF LGGVLNLNPL ERWTPQQAML HPFITKQEFT GEWFPPGSSL  
PGPSEKHDDA KGQQSEYGSA NDSSNNAGHN YVYNPSSATG GADSVDIGAI SKRKENTSGD  
ISNNFAVTHS VQEGPTSAFN KLHIVEE

## SEQ ID NO:3

GGAAACCTTC GGCCGCCGCT CCCGCCGCT ACCCGACCGA TTGGCGGCAG TAAGCACACA  
ATGAATGATC ACCTGCATGT CGGCAGCCAC GCTCACGGAC AGATCCAGGT TCAACAGTTG  
TTTGAGGATA ACAGTAACAA GCGGACAGTG CTCACGACAC AACCAAATGG GCTTACAACA  
GTGGGCAAAA CGGGCTTGCC AGTGGTGCCA GAGCGGCAGC TGGACAGCAT TCATAGACGG  
CAGGGGAGCT CCACCTCTCT AAAGTCCATG GAAGGCATGG GGAAGGTGAA AGCCACCCCC  
ATGACACCTG AACAAGCAAT GAAGCAATAC ATGCAAAAAC TCACAGCCTT CGAACACCAT  
GAGATTTTCA GCTACCCTGA AATATATTTT TTGGGTCTAA ATGCTAAGAA GCGCCAGGGC  
ATGACAGGTG GGCCCAACAA TGGTGGCTAT GATGATGACC AGGGATCATA TGTGCAGGTG  
CCCCACGATC ACGTGGCTTA CAGGTATGAG GTCCTCAAGG TCATTGGGAA GGGGAGCTTT  
GGGCAGGTGG TCAAGGCCTA CGATCACAAA GTCCACCAGC ACGTGGCCCT AAAGATGGTG  
CGGAATGAGA AGCGCTTCCA CCGCAAGCA GCGGAGGAGA TCCGAATCCT GGAACACCTG  
CGGAAGCAGG ACAAGGATAA CACAATGAAT GTCATCCATA TGCTGGAGAA TTTACCTTC  
CGCAACCACA TCTGCATGAC GTTTGAGCTG CTGAGCATGA ACCTCTATAA GCTCATCAAG  
AAGAATAAAT TCCAGGGCTT CAGTCTGCCT TTGGTTCGCA AGTTTGCCCA CTCGATTCTG  
CAGTGCTTGG ATGCTTTGCA CAAAAACAGA ATAATTCAT GTGACCTTAA GCGGAGAAC  
ATTTTGTTAA AGCAGCAGGG TAGAAGCGGT ATTAAAGTAA TTGATTTTGG CTCCAGTTGT  
TACGAGCATC AGCGTGCTA CACGTACATC CAGTCGCGTT TTTACGGGGC TCCAGAAGTG  
ATCCTTGGGG CCAGGTATGG CATGCCCAT GATATGTGGA GCCTGGGCTG CATTTTAGCA  
GAGCTCCTGA CGGGTTACCC CCTCTTGCCT GGGGAAGATG AAGGGGACCA GCTGGCCTGT  
ATGATTGAAC TGTTGGGCAT GCCCTCACAG AAACGTGCTG ATGCATCCAA ACGAGCCAAA

- 26 -

AATTTTGTGA GCTCCAAGGG TTATCCCCGT TACTGCACTG TCACGACTCT CTCAGATGGC  
TCTGTGGTCC TAAACGGAGG CCGTTCCCGG AGGGGGAAAC TGAGGGGCCC ACCGGAGAGC  
AGAGAGTGGG GTAACGCGCT GAAGGGGTGT GATGATCCCC TTTTCCTTGA CTTCTTAAAA  
CAGTGTTTAG AGTGGGATCC TGCAGTGC GC ATGACCCCAG GCCAGGCTTT GCGGCACCCC  
TGGCTGAGGA GGCGGTTGCC AAAGCCTCCC ACCGGGGAGA AAACGTCAGT GAAAAGGATA  
ACTGAGAGCA CCGGTGCTAT CACATCTATA TCCAAGTTAC CTCCACCTTC TAGCTCAGCT  
TCCAACTGA GGACTAATTT GGCGCAGATG ACAGATGCCA ATGGGAATAT TCAGCAGAGG  
ACAGTGTTGC CAAAACCTGT TAGCTGAGCT CACGTCCCCT GATGCTGGTA ACCTGAAAGA  
TACGACATTG CTGAGCCTTA CTGGGTGAA AAGGAGTAGC TCAGACCTGT TTTTATTTGC  
TCAATAACTC TACTCATTTG TATCTTTTCA GCACTTAATT TTAATGTAAG AAAGTTGTTC  
ATTTTGT TTTTATACTT TCAGAAACTT  
TTTGTGTTCT AAAAGTACAA TGAGCCTTAC TGTATTTAGT GTGGCAGAAT AATAACATCA  
ATGGCAGGCC ACTGATTACT TCATGACTGC CACGCATTTA CAGATTGGTG TCAAAGACAT  
TCACTATGTT TTTATGGTTC ATGTTATATC CTCCCCAGGG TGACAGCCCC TTAAGGCCCT  
CCTTTTCCCT CCATGCTCCA GGTCCATGCA CAGGTGTAGC ATGTC

## SEQ ID NO:4

Met Asn Asp His Leu His Val Gly Ser His Ala His Gly Gln Ile Gln Val Gln  
Gln Leu Phe Glu Asp Asn Ser Asn Lys Arg Thr Val Leu Thr Thr Gln Pro Asn  
Gly Leu Thr Thr Val Gly Lys Thr Gly Leu Pro Val Val Pro Glu Arg Gln Leu  
Asp Ser Ile His Arg Arg Gln Gly Ser Ser Thr Ser Leu Lys Ser Met Glu Gly  
Met Gly Lys Val Lys Ala Thr Pro Met Thr Pro Glu Gln Ala Met Lys Gln Tyr  
Met Gln Lys Leu Thr Ala Phe Glu His His Glu Ile Phe Ser Tyr Pro Glu Ile  
Tyr Phe Leu Gly Leu Asn Ala Lys Lys Arg Gln Gly Met Thr Gly Gly Pro Asn  
Asn Gly Gly Tyr Asp Asp Asp Gln Gly Ser Tyr Val Gln Val Pro His Asp His  
Val Ala Tyr Arg Tyr Glu Val Leu Lys Val Ile Gly Lys Gly Ser Phe Gly Gln  
Val Val Lys Ala Tyr Asp His Lys Val His Gln His Val Ala Leu Lys Met Val  
Arg Asn Glu Lys Arg Phe His Arg Gln Ala Ala Glu Glu Ile Arg Ile Leu Glu  
His Leu Arg Lys Gln Asp Lys Asp Asn Thr Met Asn Val Ile His Met Leu Glu  
Asn Phe Thr Phe Arg Asn His Ile Cys Met Thr Phe Glu Leu Leu Ser Met Asn  
Leu Tyr Lys Leu Ile Lys Lys Asn Lys Phe Gln Gly Phe Ser Leu Pro Leu Val  
Arg Lys Phe Ala His Ser Ile Leu Gln Cys Leu Asp Ala Leu His Lys Asn Arg  
Ile Ile His Cys Asp Leu Lys Pro Glu Asn Ile Leu Leu Lys Gln Gln Gly Arg  
Ser Gly Ile Lys Val Ile Asp Phe Gly Ser Ser Cys Tyr Glu His Gln Arg Val  
Tyr Thr Tyr Ile Gln Ser Arg Phe Tyr Arg Ala Pro Glu Val Ile Leu Gly Ala  
Arg Tyr Gly Met Pro Ile Asp Met Trp Ser Leu Gly Cys Ile Leu Ala Glu Leu  
Leu Thr Gly Tyr Pro Leu Leu Pro Gly Glu Asp Glu Gly Asp Gln Leu Ala Cys  
Met Ile Glu Leu Leu Gly Met Pro Ser Gln Lys Leu Leu Asp Ala Ser Lys Arg  
Ala Lys Asn Phe Val Ser Ser Lys Gly Tyr Pro Arg Tyr Cys Thr Val Thr Thr  
Leu Ser Asp Gly Ser Val Val Leu Asn Gly Gly Arg Ser Arg Arg Gly Lys Leu  
Arg Gly Pro Pro Glu Ser Arg Glu Trp Gly Asn Ala Leu Lys Gly Cys Asp Asp

- 27 -

Pro Leu Phe Leu Asp Phe Leu Lys Gln Cys Leu Glu Trp Asp Pro Ala Val Arg  
 Met Thr Pro Gly Gln Ala Leu Arg His Pro Trp Leu Arg Arg Arg Leu Pro Lys  
 Pro Pro Thr Gly Glu Lys Thr Ser Val Lys Arg Ile Thr Glu Ser Thr Gly Ala  
 Ile Thr Ser Ile Ser Lys Leu Pro Pro Pro Ser Ser Ser Ala Ser Lys Leu Arg  
 Thr Asn Leu Ala Gln Met Thr Asp Ala Asn Gly Asn Ile Gln Gln Arg Thr Val  
 Leu Pro Lys Leu Val Ser

## SEQ ID NO:5

GGAGCGAAAT GCGCTGAGCT GCAGTGTCTG GTCGAGAGTA CCCGTGGGAG CGTCGCGCCG  
 CGGAGGCAGC CGTCCCGGCG TAGGTGGCGT GGCCGACCGG ACCCCCAACT GGCGCCTCTC  
 CCCGCGCGGG GTCCCGAGCT AGGAGATGGG AGGCACAGCT CGTGGGCCTG GGCGGAAGGA  
 TGCGGGGCCG CCTGGGGCCG GGCTCCCGCC CCAGCAGCGG AGGTTGGGGG ATGGTGTCTA  
 5 TGACACCTTC ATGATGATAG ATGAAACCAA ATGTCCCCC TGTTCAAATG TACTCTGCAA  
 TCCTTCTGAA CCACCTCCAC CCAGAAGACT AAATATGACC ACTGAGCAGT TTACAGGAGA  
 TCATACTCAG CACTTTTTTG ATGGAGGTGA GATGAAGGTA GAACAGCTGT TTCAAGAATT  
 TGGCAACAGA AAATCCAATA CTATTCAGTC AGATGGCATC AGTGAAGTCTG AAAAATGCTC  
 TCCTACTGTT TCTCAGGGTA AAAGTTCAGA TTGCTTGAAT ACAGTAAAT CCAACAGTTC  
 10 ATCCAAGGCA CCCAAAGTGG TGCCTCTGAC TCCAGAACAA GCCCTGAAGC AATATAAACA  
 CCACCTCACT GCCTATGAGA AACTGGAAAT AATTAATTAT CCAGAAATTT ACTTTGTAGG  
 TCCAAATGCC AAGAAAAGAC ATGGAGTTAT TGGTGGTCCC AATAATGGAG GGTATGATGA  
 TGCAGATGGG GCCTATATTC ATGTACCTCG AGACCATCTA GCTTATCGAT ATGAGGTGCT  
 GAAAATTATT GGCAAGGGGA GTTTTGGGCA GGTGGCCAGG GTCTATGATC ACAAACCTCG  
 15 ACAGTACGTG GCCCTAAAAA TGGTGCAGCA TGAGAAGCGC TTTCATCGTC AAGCAGCTGA  
 GGAGATCCCG ATTTTGGAGC ATCTTAAGAA ACAGGATAAA ACTGGTAGTA TGAACGTTAT  
 CCACATGCTG GAAAGTTTCA CATTCCGGAA CCATGTTTGC ATGGCCTTTG AATTGCTGAG  
 CATAGACCTT TATGAGCTGA TTAACAAAAA TAAGTTTCAG GGTTTTAGCG TCCAGTTGGT  
 ACGCAAGTTT GCCCAGTCCA TCTTGCAATC TTTGGATGCC CTCCACAAAA ATAAGATTAT  
 20 TCACTGCGAT CTGAAGCCAG AAAACATTCT CCTGAAACAC CACGGGCGCA GTTCAACCAA  
 GGTCATTGAC TTTGGGTCCA GCTGTTTCGA GTACCAGAAG CTCTACACATATATCCAGTC  
 TCGGTTCTAC AGAGCTCCAG AAATCATCTT AGGAAGCCGC TACAGCACAC CAATTGACAT  
 ATGGAGTTTT GGCTGCATCC TTGCAGAACT TTAAACAGGA CAGCCTCTCT TCCCTGGAGA  
 GGATGAAGGA GACCAGTTGG CCTGCATGAT GGAGCTTCTA GGGATGCCAC CACCAAACT  
 25 TCTGGAGCAA TCCAAACGTG CCAAGTACTT TATTAATTCC AAGGGCATA CCGGCTACTG  
 CTCTGTGACT ACCCAGGCAG ATGGGAGGGT TGTGCTTGTG GGGGGTCGCT CACGTAGGGG  
 TAAAAAGCGG GGTCCCCCAG GCAGCAAAGA CTGGGGGACA GCACTGAAAG GGTGTGATGA  
 CTACTTGTTT ATAGAGTTCT TGAAAAGGTG TCTTCACTGG GACCCCTCTG CCCGCTTGAC  
 CCCAGCTCAA GCATTAAGAC ACCCTTGGAT TAGCAAGTCT GTCCCCAGAC CTCTCACCAC  
 30 CATAGACAAG GTGTCAGGGA AACGGGTAGT TAATCCTGCA AGTGCTTTCC AGGGATTGGG  
 TTCTAAGCTG CCTCCAGTTG TTGGAATAGC CAATAAGCTT AAAGCTAACT TAATGTCAGA  
 AACCAATGGT AGTATACCCC TATGCAGTGT ATTGCCAAAA CTGATTAGCT AGTGGACAGA  
 GATATGCCCA GAGATGCATA TGTGTATATT TTTATGATCT TACAAACCTG CAAATGGAAA

- 28 -

AAATGCAAGC CCATTGGTGG ATGTTTTTGT TAGAGTAGAC TTTTTTTAAA CAAGACAAAA  
CATTTTTATA TGATTATAAA A

## SEQ ID NO:6

5 MGGTARGPGR KDAGPPGAGL PPQRRRLGDG VYDTFMMIDE TKCPPCSNVL CNPSEPPPPR  
RLNMTTEQFT GDHTQHFLDG GEMKVEQLFQ EFGNRKSNTI QSDGISDSEK CSPTVSQGKS  
SDCLNTVKSN SSSKAPKVVPLTPEQALKQY KHHLTAYEKL EIINYPEIYF VGPNAKKRHG  
VIGGPNNGGY DDADGAYIHV PRDHLAYRYE VLKIIGKGSF GQVARVYDHK LRQYVALKMV  
RNEKRFHRQA AEEIRILEHL KKQDKTGSMN VIHMLSESFTF RNHVCMFEL LSIDLYELIK  
KNKFQGFVSQ LVRKFAQSIL QSLDALHKNK IIHCDLKPEN ILLKHHGRSS TKVIDFGSSC  
10 FEYQKLYTYI QSRFYRAPEI ILGSRYSTPI DIWSFGCILA ELLTGQPLFP GEDEGDQLAC  
MMELLGMPPP KLLEQSKRAK YFINSKGIPR YCSVTTQADG RVVLVGGRSR RGKKRGPPGS  
KDWGTALKGC DDYLFIEFLK RCLHWDPSAR LTPAQALRHP WISKSVPRPL TTIDKVSQGR  
VVNPASAFQG LGSKLPPVVG IANKLKANLM SETNGSIPLC SVLPKLIS

## SEQ ID NO:7

CGGCGCTGGC AAGCGAAGCT TGGGGGTGGG GAGGTAGAGT GAGCCCTCAG TAGGAGGGAC  
GAGGGCAGGG GTCTGACTGC CTCCCCGGA CCGCCCCAC CTCCTCTCTA TCAGGGCCCC  
CTCCCCCAT CCCTGTCTCA CCGGGCGCGG GGGACGGGGC TAGAGCGGAG TTAGAGCAAG  
AAGAATTTCC ACCCTGGAT TCCCTCTGAA ACCCTAGATC GGGGTATATG TTAAGGGATT  
ACGAAAATCT AGGACTTTTT GTGGGGCTTT TTATTAAAGG GGGGGAGCCC GGGAGCAATA  
CCTTGAAAG AAGCCCTGTT GCTTAGAGCG GATAACCAAC GGCTGAACTC TTGGGGTTTG  
CTGTGAGGGG TGCGGTCTAG CTTCGAATGT ACAGTGGTGG AGCCACAGTG TTAAGAACA  
GAGAAGTGAT CCTTAATCAT TTAGAATTTT GCCTCCACCA TCCACCAGAA AATGAAGTGG  
AAAGAGAAAGT TGGGGGATGG TGTCTATGAC ACCTTCATGA TGATAGATGA AACCAAATGT  
CCCCCTGTT CAAATGTACT CTGCAATCCT TCTGAACCAC CTCCACCCAG AAGACTAAAT  
ATGACCACTG AGCAGTTTAC AGGAGATCAT ACTCAGCACT TTTTGGATGG AGGTGAGATG  
AAGGTAGAAC AGCTGTTTCA AGAATTTGGC AACAGAAAAT CCAATACTAT TCAGTCAGAT  
GGCATCAGTG ACTCTGAAAA ATGCTCTCCT ACTGTTTCTC AGGGTAAAAG TTCAGATTGC  
TTGAATACAG TAAAATCCAA CAGTTCATCC AAGGCACCCA AAGTGGTGCC TCTGACTCCA  
GAACAAGCCC TGAAGCAATA TAAACACCAC CTCACTGCCT ATGAGAACT GGAAATAATT  
AATTATCCAG AAATTTACTT TGTAGGTCCA AATGCCAAGA AAAGACATGG AGTTATTGGT  
GGTCCCAATA ATGGAGGGTA TGATGATGCA GATGGGGCCT ATATTCATGT ACCTCGAGAC  
CATCTAGCTT ATCGATATGA GGTGCTGAAA ATTATTGGCA AGGGGAGTTT TGGGCAGGTG  
GCCAGGTCT ATGATCACAA ACTTCGACAG TACGTGGCCC TAAAAATGGT GCGCAATGAG  
AAGCGCTTTC ATCGTCAAGC AGCTGAGGAG ATCCGGATTT TGGAGCATCT TAAGAAACAG  
GATAAACTG GTAGTATGAA CGTTATCCAC ATGCTGGAAA GTTTCACATT CCGGAACCAT  
GTTTGATGG CCTTTGAATT GCTGAGCATA GACCTTTATG AGCTGATTAA AAAAAATAAG  
TTTCAGGGT TTAGCGTCCA GTTGGTACGC AAGTTTGCCC AGTCCATCTT GCAATCTTTG

- 29 -

GATGCCCTCC ACAAAAATAA GATTATTCAC TGCGATCTGA AGCCAGAAAA CATTCTCCTG  
AAACACCACG GGCGCAGTTC AACCAAGGTC ATTGACTTTG GGTCCAGCTG TTTCGAGTAC  
CAGAAGCTCT ACACATATAT CCAGTCTCGG TTCTACAGAG CTCCAGAAAT CATCTTAGGA  
AGCCGCTACA GCACACCAAT TGACATATGG AGTTTTGGCT GCATCCTTGC AGAACTTTTA  
ACAGGACAGC CTCTCTTCCC TGGAGAGGAT GAAGGAGACC AGTTGGCCTC CATGATGGAG  
CTTCTAGGGA TGCCACCACC AAAACTTCTG GAGCAATCCA AACGTGCCAA GTACTTTATT  
AATTCCAAGG GCATACCCCG CTACTGCTCT GTGACTACCC AGGCAGATGG GAGGGTTGTG  
CTTGTGGGGG GTCGCTCACG TAGGGGTAAA AAGCGGGGTC CCCCAGGCAG CAAAGACTGG  
GGGACAGCAC TGAAAGGGTG TGATGACTAC TTGTTTATAG AGTTCTTGAA AAGGTGTCTT  
CACTGGGACC CCTCTGCCCC CTTGACCCCA GCTCAAGCAT TAAGACACCC TTGGATTAGC  
AAGTCTGTCC CCAGACCTCT CACCACCATA GACAAGGTGT CAGGGAAACG GGTAGTTAAT  
CCTGCAAGTG CTTTCCAGGG ATTGGGTTCT AAGCTGCCTC CAGTTGTTGG AATAGCCAAT  
AAGCTTAAAG CTAACCTAAT GTCAGAAACC AATGGTAGTA TACCCCTATG CAGTGTATTG  
CCAAACTGA TTAGCTAGTG GACAGAGATA TGCCAGAGA TGCATATGTG TATATTTTAA  
TGATCTTACA AACCTGCAA TGGAAAAAT GCAAGCCCAT TGGTGGATGT TTTTGTTAGA  
GTAGACTTTT TTAAACAAG ACAAACATT TTTATATGAT TATAAAA

## SEQ ID NO:8

Met Lys Trp Lys Glu Lys Leu Gly Asp Gly Val Tyr Asp Thr Phe Met Met Ile  
Asp Glu Thr Lys Cys Pro Pro Cys Ser Asn Val Leu Cys Asn Pro Ser Glu Pro  
Pro Pro Pro Arg Arg Leu Asn Met Thr Thr Glu Gln Phe Thr Gly Asp His Thr  
Gln His Phe Leu Asp Gly Gly Glu Met Lys Val Glu Gln Leu Phe Gln Glu Phe  
Gly Asn Arg Lys Ser Asn Thr Ile Gln Ser Asp Gly Ile Ser Asp Ser Glu Lys  
Cys Ser Pro Thr Val Ser Gln Gly Lys Ser Ser Asp Cys Leu Asn Thr Val Lys  
Ser Asn Ser Ser Ser Lys Ala Pro Lys Val Val Pro Leu Thr Pro Glu Gln Ala  
Leu Lys Gln Tyr Lys His His Leu Thr Ala Tyr Glu Lys Leu Glu Ile Ile Asn  
Tyr Pro Glu Ile Tyr Phe Val Gly Pro Asn Ala Lys Lys Arg His Gly Val Ile  
Gly Gly Pro Asn Asn Gly Gly Tyr Asp Asp Ala Asp Gly Ala Tyr Ile His Val  
Pro Arg Asp His Leu Ala Tyr Arg Tyr Glu Val Leu Lys Ile Ile Gly Lys Gly  
Ser Phe Gly Gln Val Ala Arg Val Tyr Asp His Lys Leu Arg Gln Tyr Val Ala  
Leu Lys Met Val Arg Asn Glu Lys Arg Phe His Arg Gln Ala Ala Glu Glu Ile  
Arg Ile Leu Glu His Leu Lys Lys Gln Asp Lys Thr Gly Ser Met Asn Val Ile  
His Met Leu Glu Ser Phe Thr Phe Arg Asn His Val Cys Met Ala Phe Glu Leu  
Leu Ser Ile Asp Leu Tyr Glu Leu Ile Lys Lys Asn Lys Phe Gln Gly Phe Ser  
Val Gln Leu Val Arg Lys Phe Ala Gln Ser Ile Leu Gln Ser Leu Asp Ala Leu  
His Lys Asn Lys Ile Ile His Cys Asp Leu Lys Pro Glu Asn Ile Leu Leu Lys  
His His Gly Arg Ser Ser Thr Lys Val Ile Asp Phe Gly Ser Ser Cys Phe Glu  
Tyr Gln Lys Leu Tyr Thr Tyr Ile Gln Ser Arg Phe Tyr Arg Ala Pro Glu Ile  
Ile Leu Gly Ser Arg Tyr Ser Thr Pro Ile Asp Ile Trp Ser Phe Gly Cys Ile  
Leu Ala Glu Leu Leu Thr Gly Gln Pro Leu Phe Pro Gly Glu Asp Glu Gly Asp  
Gln Leu Ala Ser Met Met Glu Leu Leu Gly Met Pro Pro Pro Lys Leu Leu Glu

- 30 -

Gln Ser Lys Arg Ala Lys Tyr Phe Ile Asn Ser Lys Gly Ile Pro Arg Tyr Cys  
Ser Val Thr Thr Gln Ala Asp Gly Arg Val Val Leu Val Gly Gly Arg Ser Arg  
Arg Gly Lys Lys Arg Gly Pro Pro Gly Ser Lys Asp Trp Gly Thr Ala Leu Lys  
Gly Cys Asp Asp Tyr Leu Phe Ile Glu Phe Leu Lys Arg Cys Leu His Trp Asp  
Pro Ser Ala Arg Leu Thr Pro Ala Gln Ala Leu Arg His Pro Trp Ile Ser Lys  
Ser Val Pro Arg Pro Leu Thr Thr Ile Asp Lys Val Ser Gly Lys Arg Val Val  
Asn Pro Ala Ser Ala Phe Gln Gly Leu Gly Ser Lys Leu Pro Pro Val Val Gly  
Ile Ala Asn Lys Leu Lys Ala Asn Leu Met Ser Glu Thr Asn Gly Ser Ile Pro  
Leu Cys Ser Val Leu Pro Lys Leu Ile Ser

**SEQ ID NO:9**

Leu Gly Gly Arg Asp Ser Arg Ser Gly Ser Pro Met Ala Arg Arg

**SEQ ID NO:10**

Leu Gly Gly Arg Asp Ser Arg Ser Gly Ser Ala Met Ala Arg Arg (Peptide 1)

**5 SEQ ID NO:11**

Leu Gly Gly Arg Asp Ser Arg Ser Gly Thr Pro Met Ala Arg Arg (Peptide 2)

**SEQ ID NO:12**

Leu Gly Gly Arg Asp Ser Arg Ser Gly Ser Pro Pro Ala Arg Arg (Peptide 3)

**SEQ ID NO:13**

10 Leu Gly Gly Arg Asp Ser Arg Ser Pro Ser Pro Met Ala Arg Arg (Peptide 4)

**SEQ ID NO:14**

Leu Gly Gly Arg Asp Ser Arg Ser Pro Ser Ala Met Ala Arg Arg (Peptide 5)

**SEQ ID NO:15**

Leu Gly Gly Arg Asp Ser Arg Ala Gly Ser Pro Met Ala Arg Arg (peptide 6)

**15 SEQ ID NO:16**

Leu Gly Gly Arg Asp Ala Arg Ser Gly Ser Pro Met Ala Arg Arg (Peptide 7)

**SEQ ID NO:17**

- 31 -

Leu Gly Gly Arg Asp Ala Arg Ala Gly Ser Pro Met Ala Arg Arg (peptide 8)

**SEQ ID NO:18**

NIVTPRTPPPSQGK



**What is claimed is:**

1. A method for screening for an agent that modulates a YAK polypeptide kinase activity comprising:

5 a. contacting a candidate agent with a YAK polypeptide, wherein the step of contacting is carried out under conditions and for a time sufficient to allow the candidate agent and the polypeptide to interact; and

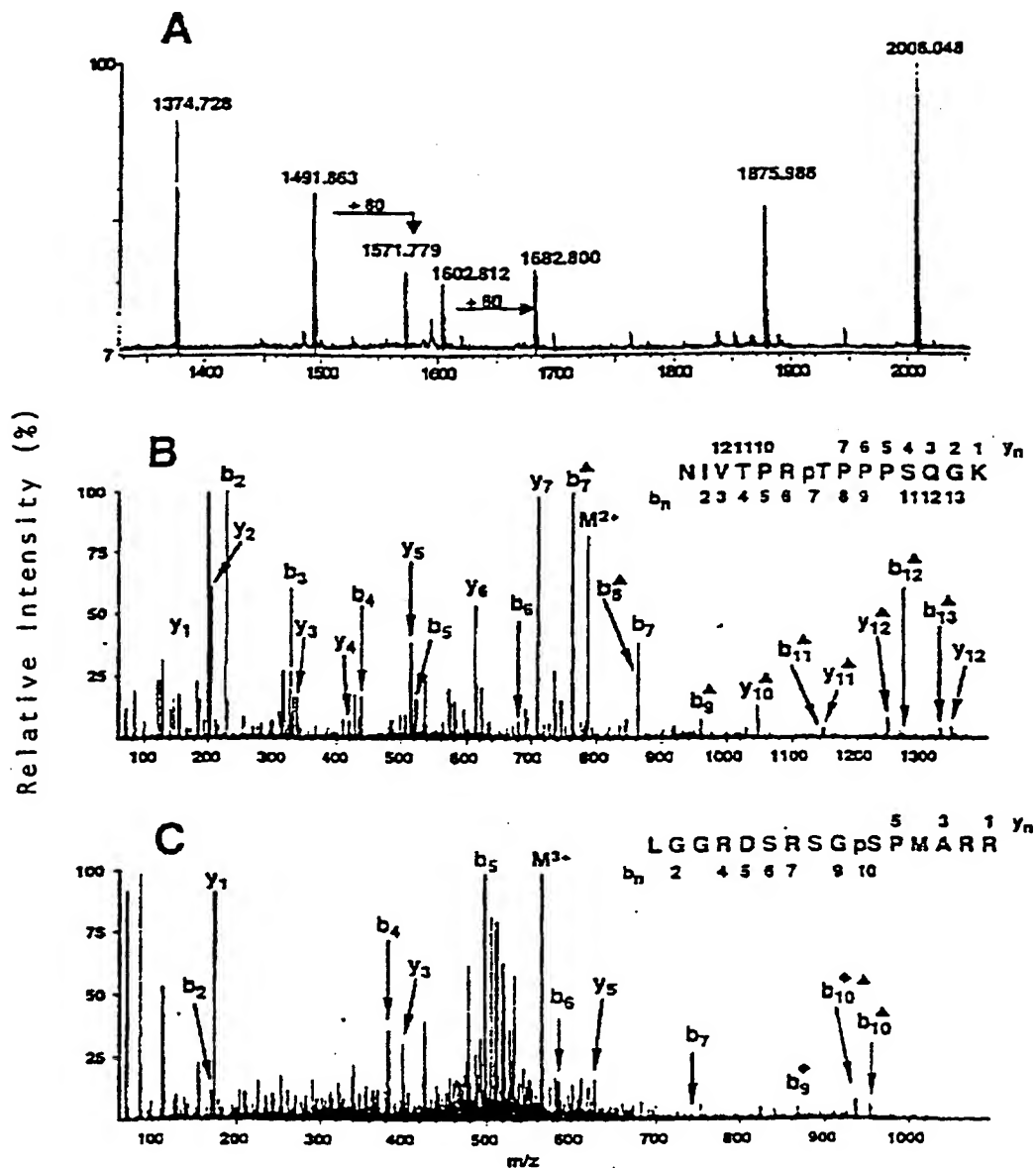
b. subsequently measuring the ability of the candidate agent to modulate the ability of the polypeptide to phosphorylate the substrate.

10 2. The method of claim 1 wherein the YAK polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, and SEQ ID NO:8.

3. The method of claim 2 wherein the substrate comprises the SER-164 polypeptide.

1/3

Figure 1

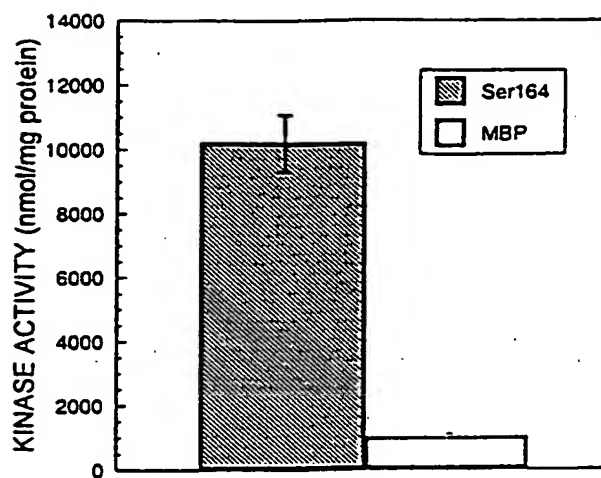


BEST AVAILABLE COPY

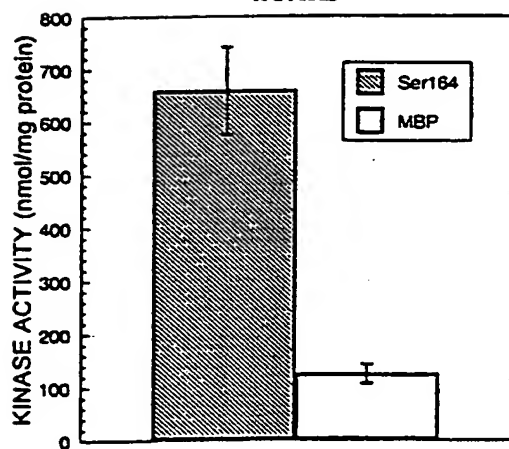
2/3

Figure 2

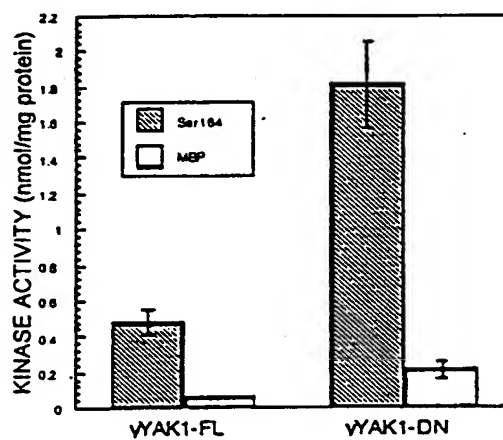
## hYAK1



## hYAK3



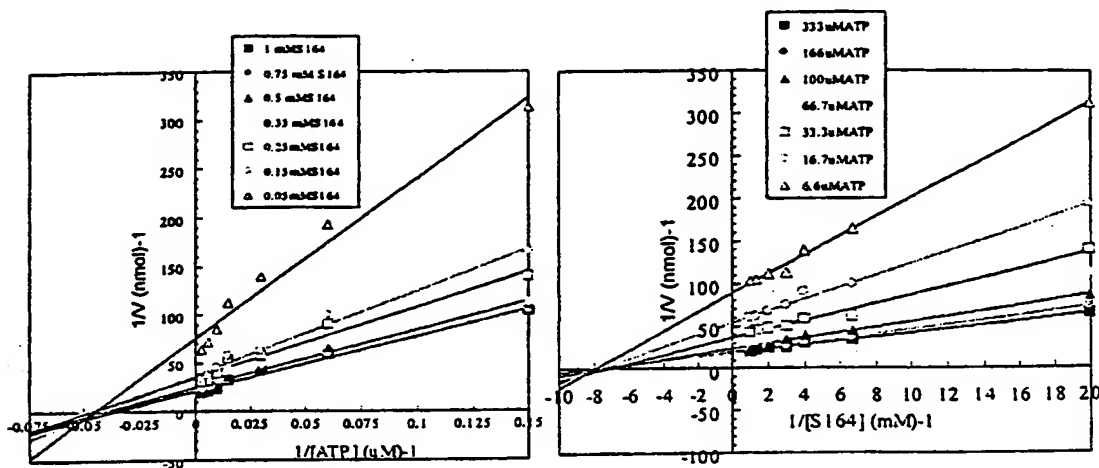
## Yeast YAK1



BEST AVAILABLE COPY

Figure 3

## Steady State Two Substrate Kinetic Analysis

Kinetic Constants (GraphFit) :

$$K_m[\text{ATP}] = 42 \pm 7 \text{ } \mu\text{M} \quad K_m[\text{S164}] = 160 \pm 14 \text{ } \mu\text{M}$$

$$V_{max} = 51 \pm 6 \text{ } \mu\text{mol/mg} \quad k_{cat} = 160 \pm 19 \text{ min}^{-1}$$

## 1.

## SEQUENCE LISTING

<110> SMITHKLINE BEECHAM CORPORATION  
SMITHKLINE BEEHCAM PLC

<120> SPECIFIC PEPTIDE SUBSTRATE FOR HUMAN  
YAK1, YAK3A, AND YAK3B AND YEAST YAK1 PROTEIN KINASES

<130> GH-70650WO

<140> TO BE ASSIGNED

<141> 2000-11-06

<150> 60/163,901

<151> 1999-11-05

<150> TO BE ASSIGNED \_\_\_\_\_

<151> 2000-11-03

<160> 18

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 2424

<212> DNA

<213> HOMO SAPIENS

<400> 1

atgaactcat	ccaataataa	cgactcgtcc	agctccaata	gcaacatgaa	taactccttg	60
agcccgaccc	ttgtgaccca	cagtgatgct	agtatgggct	cgggtagagc	aagcccagac	120
aatagccata	tggggagagg	tatatggaat	ccatcgtacg	taaatcaagg	ctctcaaagg	180
tctccacagc	agcagcatca	gaatcatcac	cagcagcagc	agcagcagca	acaacaacaa	240
caacagaatt	ctcaattttg	ctttgttaac	ccttggaatg	aggaaaaagt	aactaattct	300
caacaaaacc	tggtgtatcc	ccctcaatac	gatgacttaa	acagtaacga	aagtctagat	360
gcgtacagac	gacgtaaatc	tagtctcggt	gtacctccag	ccagggcacc	tgccccaagt	420
cctttccagt	acgatagtta	tcccgtttac	accagctcta	atacagagtt	ggcaggaaat	480
agcagtggcc	agtatccttc	tggtatcaa	caacaacaac	agcaagtata	ccagcagggc	540
gctatccatc	cttccaatt	tggtatccaga	tttgttccct	ccctttatga	tcgtcaagat	600
ttccaaagaa	ggcagagtct	ggctgcaact	aattattcgt	ccaatttttc	ttccctaaat	660
tcaaatacta	atcaaggaac	aaactccata	ccagtcatgt	caccttacag	gaggcttagc	720
gcatacctc	ctagtacaag	tccccacta	caacccccct	tcaagcagtt	acgaagagat	780
gaagtacaag	gtcagaaatt	gtccattcct	cagatgcaac	tttgcaactc	taaaaacgat	840
cttcaacctg	tcttaaatgc	tactccaaag	tttagacgtg	cgtcgctcaa	ttccaagacg	900
atatctcccc	tagtcagtgt	cacgaaaagc	ctgattacca	catattcttt	atgttctcca	960
gaattttacct	accaaagctc	caaaaaaccc	aagagagtac	ttacgaagcc	cagtgaaggg	1020
aaatgtaaca	acggattcga	caacataaac	agtgactata	ttctttatgt	aaatgacgtt	1080
ttgggtgtag	agcagaacag	aaagtacctt	gtgctagaca	ttttgggtca	aggtacattt	1140
ggtcgaagtgg	tcaaatgtca	aaatttgctg	acgaaagaga	tattggctgt	aaaagtgggt	1200
aaatcgagga	cagaatattt	gactcaaagt	ataacggagg	ctaaaatttt	agagctactg	1260
aatcaaaaaga	tagaccctac	taataaacat	catttttttaa	ggatgtatga	ctcctttgtc	1320
cacaagaacc	acttatgttt	agtgtttgaa	ttactaagca	ataatttgta	cgaattatta	1380

```

aaacaaaaca aatttcattg gctctctata caacttatca gaacattcac cactcaaata 1440
ctggattcgt tatgcgtttt gaaggaaagc aaactgattc attgcgatct gaagccagag 1500
aatatcttgc tctgcgcacc tgataagccg gaattgaaaa ttattgattt tggctcatct 1560
tgcaagagg caagaaccgt ttatacatat atccagtcca ggttttatcg tgctcctgaa 1620
attatactgg gtataaccgt ttcaaccagt attgacatgt ggtcggttagg ttgtattggt 1680
gctgaattat ttttgggtat accgatcttc ccaggcgctt ctgaatataa ccaattaaca 1740
agaataatag acacgcttgg atatcctcca tcgtggatga tagatatggg taaaaactct 1800
ggaaaattta tgaagaaatt ggcaccagaa gaaagttctt cttctacaca aaagcatcgt 1860
atgaaaacta ttgaagagtt ttgcagagaa tacaatatag tggaaaagcc cagtaaacaa 1920
tattttaagt ggagaaagtt accagatatt attagaaact acaggtatcc taaaagcata 1980
cagaactccc aagaacttat cgaccaagaa atgcagaata gggagtgttt gatccacttt 2040
ttaggcggtg tgctaaattt gaaccggtta gaaagatgga caccacaaca agctatgcta 2100
cacccttca taacaaagca ggagtttaca ggtgagtggg ttcctccagg atcgtcttta 2160
ccgggtcctt cagaaaaaca tgacgatgca aaaggccagc aaagtgaata tggaaagtgcg 2220
aacgactcta gtaacaatgc aggccacaac tatgtctata atcctagctc tgccactggt 2280
gggtgctgata gcgtcgacat tgggtgctatc agtaaaagga aggagaatac atctggcgac 2340
atctccaata attttgctgt tactcattct gttcaagaag ggccaacaag cgcgttcaat 2400
aaacttcaca ttgtcgaaga ataa 2424

```

&lt;210&gt; 2

&lt;211&gt; 807

&lt;212&gt; PRT

&lt;213&gt; HOMO SAPIENS

&lt;400&gt; 2

```

Met Asn Ser Ser Asn Asn Asn Asp Ser Ser Ser Ser Asn Ser Asn Met
1 5 10 15
Asn Asn Ser Leu Ser Pro Thr Leu Val Thr His Ser Asp Ala Ser Met
20 25 30
Gly Ser Gly Arg Ala Ser Pro Asp Asn Ser His Met Gly Arg Gly Ile
35 40 45
Trp Asn Pro Ser Tyr Val Asn Gln Gly Ser Gln Arg Ser Pro Gln Gln
50 55 60
Gln His Gln Asn His His Gln Gln Gln Gln Gln Gln Gln Gln Gln
65 70 75 80
Gln Gln Asn Ser Gln Phe Cys Phe Val Asn Pro Trp Asn Glu Glu Lys
85 90 95
Val Thr Asn Ser Gln Gln Asn Leu Val Tyr Pro Pro Gln Tyr Asp Asp
100 105 110
Leu Asn Ser Asn Glu Ser Leu Asp Ala Tyr Arg Arg Arg Lys Ser Ser
115 120 125
Leu Val Val Pro Pro Ala Arg Ala Pro Ala Pro Asn Pro Phe Gln Tyr
130 135 140
Asp Ser Tyr Pro Ala Tyr Thr Ser Ser Asn Thr Ser Leu Ala Gly Asn
145 150 155 160
Ser Ser Gly Gln Tyr Pro Ser Gly Tyr Gln Gln Gln Gln Gln Gln Val
165 170 175
Tyr Gln Gln Gly Ala Ile His Pro Ser Gln Phe Gly Ser Arg Phe Val
180 185 190
Pro Ser Leu Tyr Asp Arg Gln Asp Phe Gln Arg Arg Gln Ser Leu Ala
195 200 205
Ala Thr Asn Tyr Ser Ser Asn Phe Ser Ser Leu Asn Ser Asn Thr Asn
210 215 220
Gln Gly Thr Asn Ser Ile Pro Val Met Ser Pro Tyr Arg Arg Leu Ser

```



```
<210> 3
<211> 2085
<212> DNA
<213> HOMO SAPIENS
```

<400> 3						
ggaaaccttc	ggccgcgcgt	ccgcgcgcct	ccccgaccga	ttggcggcag	taagcacaca	60
atgaatgatc	acctgcatgt	cggcagccac	gctcacggac	agatccaggt	tcaacagttg	120
tttgaggata	acagtaacaa	gcggacagtg	ctcacgacac	aaccaaattg	gcttacaaca	180
gtggggcaaaa	cgggcttgcc	agtggtgcca	gagcggcagc	tggacagcat	tcatgacggg	240
caggggagct	ccacctctct	aaagtcctat	gaaggcatgg	ggaagtgtaa	acgacacccc	300
atgacacctg	acaagcaaat	gaagcaatac	atgcaaaaaa	tcacagcctt	cgaacaccat	360
gagattttca	gctaccctga	aatatatattc	ttgggtctaa	atgctaagaa	gcgccagggc	420
atgacaggtg	ggcccaacaa	tgggtggctat	gatgatgacc	agggatcata	tgtgcaggtg	480
ccccacgatc	acgtggctta	caggtatgag	gtcctcaagg	tcattgggaa	ggggagcttt	540
gggcaggtgg	tcaaggccta	cgatcacaaa	gtccaccagg	acgtggccct	aaagatgggtg	600
cggaatgaga	agcgcttcca	ccggcaagca	gcggaggaga	tccgaatcct	ggaacacctg	660
cggaagcagg	acaaggataa	cacaatgaat	gtcatccata	tgtctggagaa	tttcaccttc	720
cgcaaccaca	ctctgcatgac	gttttagactg	ctgagcatga	acctctataa	gctcatcaag	780
aagaataaat	tccagggttt	cagtctgctt	ttggttcgca	agtttgccca	ctcgattctg	840
cagtgccttg	atgctttgca	caaaaacaga	ataattcact	gtgaccttaa	gcccgagaac	900
attttgttaa	agcagcaggg	tagaagcgg	attaaagtaa	ttgattttgg	ctccagttgt	960
tacgagcatc	agcgtgtcta	cacgtacatc	cagtccgctt	tttaccgggc	tccagaagtg	1020
atccttgggg	ccagggtatg	catgcccatt	gatatgtgga	gcttgggtgc	catttttagca	1080
gagctcctga	cgggttaacc	cctcttgctt	ggggaagatg	aaggggacca	gctggcctgt	1140
atgattgaac	tgttgggcgt	gccttcacag	aaactgctgg	atgcattcaa	acgagccaaa	1200
aattttgtga	gttccaaggg	ttatccccgt	tactgcactg	tcacgactct	ctcagatggc	1260
tctgtggtcc	taaacggagg	ccgttcccgg	agggggaaac	tgaggggcc	accggagagc	1320
agagagtggg	gtaacgcgct	gaaggggtgt	gatgatcccc	ttttccttga	cttcttaaaa	1380
cagtgtttag	agtgggatcc	tgacgtgcgc	atgacccag	gccaggcttt	gcggcacccc	1440
tggctgagga	ggcggttgcc	aaagcctccc	accggggaga	aaacgtcagt	gaaaaggata	1500
actgagagca	cgggtgctat	cacatctata	tccaagttac	ctccaccttc	tagctcagct	1560
tccaaactga	ggactaattt	ggcgcagatg	acagatgcca	atgggaatat	tcagcagagg	1620
acagtgttgc	caaaacttgt	tagctgagct	cacgtcccct	gatgctggta	acctgaaaga	1680



```

tacgacattg ctgagcctta ctgggttgaa aaggagtagc tcagacctgt ttttatttgc 1740
tcaataactc tactcatttg tatcttttca gcacttaatt ttaatgtaag aaagttgttc 1800
atattgtttt tataaaatac atgaggacaa tgctttaagt ttttatactt tcagaaactt 1860
tttgtgttct aaaagtacaa tgagccttac tgtatttagt gtggcagaat aataacatca 1920
atggcaggcc actgattact tcatgactgc cacgcattta cagattggtg tcaaagacat 1980
tcactatgtt tttatgggtc atgttatatc ctccccaggg tgacagcccc ttaaggccct 2040
ccttttccct ccatgctcca ggtccatgca caggtgtagc atgtc 2085

```

<210> 4  
 <211> 528  
 <212> PRT  
 <213> HOMO SAPIENS

<400> 4

Met	Asn	Asp	His	Leu	His	Val	Gly	Ser	His	Ala	His	Gly	Gln	Ile	Gln
1				5					10					15	
Val	Gln	Gln	Leu	Phe	Glu	Asp	Asn	Ser	Asn	Lys	Arg	Thr	Val	Leu	Thr
			20					25					30		
Thr	Gln	Pro	Asn	Gly	Leu	Thr	Thr	Val	Gly	Lys	Thr	Gly	Leu	Pro	Val
		35					40					45			
Val	Pro	Glu	Arg	Gln	Leu	Asp	Ser	Ile	His	Arg	Arg	Gln	Gly	Ser	Ser
	50					55					60				
Thr	Ser	Leu	Lys	Ser	Met	Glu	Gly	Met	Gly	Lys	Val	Lys	Ala	Thr	Pro
65					70					75					80
Met	Thr	Pro	Glu	Gln	Ala	Met	Lys	Gln	Tyr	Met	Gln	Lys	Leu	Thr	Ala
				85					90					95	
Phe	Glu	His	His	Glu	Ile	Phe	Ser	Tyr	Pro	Glu	Ile	Tyr	Phe	Leu	Gly
			100					105					110		
Leu	Asn	Ala	Lys	Lys	Arg	Gln	Gly	Met	Thr	Gly	Gly	Pro	Asn	Asn	Gly
		115					120					125			
Gly	Tyr	Asp	Asp	Asp	Gln	Gly	Ser	Tyr	Val	Gln	Val	Pro	His	Asp	His
	130					135						140			
Val	Ala	Tyr	Arg	Tyr	Glu	Val	Leu	Lys	Val	Ile	Gly	Lys	Gly	Ser	Phe
145					150					155					160
Gly	Gln	Val	Val	Lys	Ala	Tyr	Asp	His	Lys	Val	His	Gln	His	Val	Ala
			165					170						175	
Leu	Lys	Met	Val	Arg	Asn	Glu	Lys	Arg	Phe	His	Arg	Gln	Ala	Ala	Glu
			180				185						190		
Glu	Ile	Arg	Ile	Leu	Glu	His	Leu	Arg	Lys	Gln	Asp	Lys	Asp	Asn	Thr
	195						200					205			
Met	Asn	Val	Ile	His	Met	Leu	Glu	Asn	Phe	Thr	Phe	Arg	Asn	His	Ile
	210					215					220				
Cys	Met	Thr	Phe	Glu	Leu	Leu	Ser	Met	Asn	Leu	Tyr	Lys	Leu	Ile	Lys
225					230					235					240
Lys	Asn	Lys	Phe	Gln	Gly	Phe	Ser	Leu	Pro	Leu	Val	Arg	Lys	Phe	Ala
			245						250					255	
His	Ser	Ile	Leu	Gln	Cys	Leu	Asp	Ala	Leu	His	Lys	Asn	Arg	Ile	Ile
			260				265						270		
His	Cys	Asp	Leu	Lys	Pro	Glu	Asn	Ile	Leu	Leu	Lys	Gln	Gln	Gly	Arg
		275					280					285			
Ser	Gly	Ile	Lys	Val	Ile	Asp	Phe	Gly	Ser	Ser	Cys	Tyr	Glu	His	Gln
	290					295					300				
Arg	Val	Tyr	Thr	Tyr	Ile	Gln	Ser	Arg	Phe	Tyr	Arg	Ala	Pro	Glu	Val
305					310					315					320

Ile Leu Gly Ala Arg Tyr Gly Met Pro Ile Asp Met Trp Ser Leu Gly  
 325 330 335  
 Cys Ile Leu Ala Glu Leu Leu Thr Gly Tyr Pro Leu Leu Pro Gly Glu  
 340 345 350  
 Asp Glu Gly Asp Gln Leu Ala Cys Met Ile Glu Leu Leu Gly Met Pro  
 355 360 365  
 Ser Gln Lys Leu Leu Asp Ala Ser Lys Arg Ala Lys Asn Phe Val Ser  
 370 375 380  
 Ser Lys Gly Tyr Pro Arg Tyr Cys Thr Val Thr Thr Leu Ser Asp Gly  
 385 390 395 400  
 Ser Val Val Leu Asn Gly Gly Arg Ser Arg Arg Gly Lys Leu Arg Gly  
 405 410 415  
 Pro Pro Glu Ser Arg Glu Trp Gly Asn Ala Leu Lys Gly Cys Asp Asp  
 420 425 430  
 Pro Leu Phe Leu Asp Phe Leu Lys Gln Cys Leu Glu Trp Asp Pro Ala  
 435 440 445  
 Val Arg Met Thr Pro Gly Gln Ala Leu Arg His Pro Trp Leu Arg Arg  
 450 455 460  
 Arg Leu Pro Lys Pro Pro Thr Gly Glu Lys Thr Ser Val Lys Arg Ile  
 465 470 475 480  
 Thr Glu Ser Thr Gly Ala Ile Thr Ser Ile Ser Lys Leu Pro Pro Pro  
 485 490 495  
 Ser Ser Ser Ala Ser Lys Leu Arg Thr Asn Leu Ala Gln Met Thr Asp  
 500 505 510  
 Ala Asn Gly Asn Ile Gln Gln Arg Thr Val Leu Pro Lys Leu Val Ser  
 515 520 525

&lt;210&gt; 5

&lt;211&gt; 2061

&lt;212&gt; DNA

&lt;213&gt; HOMO SAPIENS

&lt;400&gt; 5

ggagcgaaat	gcgctgagct	gcagtgtctg	gtcgagagta	cccgtgggag	cgtcgcgcgcg	60
cggaggcagc	cgtcccggcg	taggtggcgt	ggccgaccgg	acccccaaact	ggcgcctctc	120
cccgcgcggg	gtcccagact	aggagatggg	aggcacagct	cgtgggcctg	ggcgggaagga	180
tgcggggccg	cctggggccg	ggctcccgcc	ccagcagcgg	agggtggggg	atggtgtcta	240
tgacaccttc	atgatgatag	atgaaaccaa	atgtccccc	tgttcaaagt	tactctgcaa	300
tccttctgaa	ccacctccac	ccagaagact	aaatatgacc	actgagcagt	ttacaggaga	360
tcatactcag	cacttttttg	atggaggtga	gatgaaggta	gaacagctgt	ttcaagaatt	420
tggcaacaga	aaatccaata	ctattcagtc	agatggcatc	agtgactctg	aaaaatgctc	480
tcctactgtt	tctcagggtg	aaagttcaga	ttgcttgaat	acagtaaaat	ccaacagttc	540
atccaaggca	cccaaagtgg	tgctctgac	tccagaacaa	gccctgaagc	aatataaaca	600
ccacctcact	gcctatgaga	aactggaaat	aattaattat	ccagaaat	actttgtagg	660
tccaaatgcc	aagaaaagac	atggagttat	tggtgggtccc	aataatggag	ggtatgatga	720
tgcagatggg	gcctatat	atgtacctcg	agaccatcta	gcttatcgat	atgaggtgct	780
gaaaattatt	ggcaagggga	gttttgggca	ggtggccagg	gtctatgatc	acaaacttcg	840
acagtacgtg	gccctaaaaa	tggtgcgcaa	tgagaagcgc	tttcatcgtc	aagcagctga	900
ggagatccgg	attttgagc	atcttaagaa	acaggataaa	actggtagta	tgaacgttat	960
ccacatgctg	gaaagtttca	cattccggaa	ccatgtttgc	atggcctttg	aattgctgag	1020
catagacctt	tatgagctga	ttaaaaaaa	taagtttcag	ggttttagcg	tccagttggt	1080
acgcaagttt	gccagttcca	tcttgcaatc	tttggatgcc	ctccacaaaa	ataagattat	1140
tcactgcgat	ctgaagccag	aaaacattct	cctgaaacac	cacgggcgca	gttcaaccaa	1200
ggtcattgac	tttgggtcca	gctgtttcga	gtaccagaag	ctctacacat	atatccagtc	1260

```

tcggttctac agagctccag aaatcatctt aggaagccgc tacagcacac caattgacat 1320
atggagtttt ggctgcatcc ttgcagaact tttaacagga cagcctctct tccctggaga 1380
ggatgaagga gaccagttgg cctgcatgat ggagcttcta gggatgccac caccaaaact 1440
tctggagcaa tccaaacgtg ccaagtactt tattaattcc aagggcatac cccgctactg 1500
ctctgtgact acccaggcag atgggagggt tgtgcttggt gggggtcgct cacgtagggg 1560
taaaaagcgg ggtcccccag gcagcaaaga ctgggggaca gcactgaaag ggtgtgatga 1620
ctacttgttt atagagttct tgaaaagggt tcttcactgg gacccctctg cccgcttgac 1680
cccagctcaa gcattaagac acccttggat tagcaagtct gtcccagac ctctcaccac 1740
catagacaag gtgtcagggg aacgggtagt taatcctgca agtgctttcc agggattggg 1800
ttctaagctg cctccagttg ttggaatagc caataagctt aaagctaact taatgtcaga 1860
aaccaatggt agtatacccc tatgcagtgt attgccaaaa ctgattagct agtggacaga 1920
gatatgccca gagatgcata tgtgtatatt tttatgatct tacaaacctg caaatggaaa 1980
aaatgcaagc ccattgggtg atgtttttgt tagagtagac tttttttaa caagacaaaa 2040
cattttttata tgattataaa a 2061

```

```

<210> 6
<211> 588
<212> PRT
<213> HOMO SAPIENS

```

```

<400> 6
Met Gly Gly Thr Ala Arg Gly Pro Gly Arg Lys Asp Ala Gly Pro Pro
1      5      10      15
Gly Ala Gly Leu Pro Pro Gln Gln Arg Arg Leu Gly Asp Gly Val Tyr
20      25      30
Asp Thr Phe Met Met Ile Asp Glu Thr Lys Cys Pro Pro Cys Ser Asn
35      40      45
Val Leu Cys Asn Pro Ser Glu Pro Pro Pro Arg Arg Leu Asn Met
50      55      60
Thr Thr Glu Gln Phe Thr Gly Asp His Thr Gln His Phe Leu Asp Gly
65      70      75      80
Gly Glu Met Lys Val Glu Gln Leu Phe Gln Glu Phe Gly Asn Arg Lys
85      90      95
Ser Asn Thr Ile Gln Ser Asp Gly Ile Ser Asp Ser Glu Lys Cys Ser
100     105     110
Pro Thr Val Ser Gln Gly Lys Ser Ser Asp Cys Leu Asn Thr Val Lys
115     120     125
Ser Asn Ser Ser Ser Lys Ala Pro Lys Val Val Pro Leu Thr Pro Glu
130     135     140
Gln Ala Leu Lys Gln Tyr Lys His His Leu Thr Ala Tyr Glu Lys Leu
145     150     155     160
Glu Ile Ile Asn Tyr Pro Glu Ile Tyr Phe Val Gly Pro Asn Ala Lys
165     170     175
Lys Arg His Gly Val Ile Gly Gly Pro Asn Asn Gly Gly Tyr Asp Asp
180     185     190
Ala Asp Gly Ala Tyr Ile His Val Pro Arg Asp His Leu Ala Tyr Arg
195     200     205
Tyr Glu Val Leu Lys Ile Ile Gly Lys Gly Ser Phe Gly Gln Val Ala
210     215     220
Arg Val Tyr Asp His Lys Leu Arg Gln Tyr Val Ala Leu Lys Met Val
225     230     235     240
Arg Asn Glu Lys Arg Phe His Arg Gln Ala Ala Glu Glu Ile Arg Ile
245     250     255
Leu Glu His Leu Lys Lys Gln Asp Lys Thr Gly Ser Met Asn Val Ile

```

```
<210> 7
<211> 2327
<212> DNA
<213> HOMO SAPIENS
```

<400> 7						
cggcgctggc	aagcgaagct	tgggggtggg	gaggtagagt	gagccctcag	taggagggac	60
gagggcaggg	gtctgactgc	ctccccggga	ccgccccac	ctcctctcta	tcagggcccc	120
ctccccccat	ccctgtctca	cgggcgcgg	gggacgggc	tagagcggag	ttagagcaag	180
aagaatttcc	acccctggat	tccctctgaa	accctagatc	ggggtatatg	ttaagggatt	240
acgaaaatct	aggacttttt	gtggggcttt	ttattaaagg	gggggagccc	gggagcaata	300
ccttgqaaag	aagccctgtt	gcttagagcg	gataaccaac	ggctgaactc	ttgggggtttg	360

```

ctgtgagggg  tgcgggtctag  cttcgaatgt  acagtgggtg  agccacagtg  ttaaagaaca  420
gagaagtgat  ccttaatcat  ttagaatttt  gcctccacca  tccaccagaa  aatgaagtgg  480
aaagagaagt  tgggggatgg  tgtctatgac  accttcatga  tgatagatga  aaccaaatgt  540
ccccctgttt  caaatgtact  ctgcaatcct  tctgaaccac  ctccaccag  aagactaaat  600
atgaccactg  agcagtttac  aggagatcat  actcagcact  ttttgatgg  aggtgagatg  660
aaggtagaac  agctgtttca  agaatttggc  aacagaaaat  ccaatactat  tcagtcagat  720
ggcatcagtg  actctgaaaa  atgctctcct  actgtttctc  agggtaaaag  ttcagattgc  780
ttgaatacag  taaaatccaa  cagttcatcc  aaggcaccce  aagtggtgcc  tctgactcca  840
gaacaagccc  tgaagcaata  taaacaccac  ctcaactgct  atgagaaact  ggaaataatt  900
aattatccag  aaatttactt  ttaggttcca  aatgccaaag  aaagacatgg  agttattggt  960
ggtcccaata  atggagggtg  tgatgatgca  gatggggcct  atattcatgt  acctcgagac  1020
catctagctt  atcgatatga  ggtgctgaaa  attattggca  aggggagttt  tgggcaggtg  1080
gccagggtct  atgatcacia  acttcgacag  tacgtggccc  taaaaatggt  gcgcaatgag  1140
aagcgctttc  atcgtcaagc  agctgaggag  atccggattt  tggagcatct  taagaaacag  1200
gataaaactg  gtagtatgaa  cgttatccac  atgctggaaa  gtttcacatt  ccggaaccat  1260
gtttgcatgg  cctttgaatt  gctgagcata  gacctttatg  agctgattaa  aaaaaataag  1320
tttcaggggt  ttagcgtcca  gttggtacgc  aagtttgccc  agtccatctt  gcaatctttg  1380
gatgccctcc  acaaaaataa  gattattcac  tgcgatctga  agccagaaaa  cattctcctg  1440
aaacaccacg  ggcgcagttc  aaccaaggte  attgactttg  ggtccagctg  tttcgagtac  1500
cagaagctct  acacatatat  ccagtctcgg  ttctacagag  ctccagaaat  catcttagga  1560
agccgctaca  gcacaccaat  tgacatatgg  agttttggct  gcatccttgc  agaactttta  1620
acaggacagc  ctctcttccc  tggagaggat  gaaggagacc  agttggcctc  catgatggag  1680
cttctaggga  tgccaccacc  aaaacttctg  gagcaatcca  aacgtgccaa  gtactttatt  1740
aattccaagg  gcataccccg  ctactgctct  gtgactacc  aggagatgg  gagggttgtg  1800
cttggtgggg  gtcgctcacg  taggggtaaa  aagcggggtc  cccagggcag  caaagactgg  1860
gggacagcac  tgaaagggtg  tgatgactac  ttgtttatag  agttcttgaa  aaggtgtctt  1920
cactgggacc  cctctgcccg  cttagcccca  gctcaagcat  taagacacc  ttggattagc  1980
aagtctgtcc  ccagacctct  caccaccata  gacaagggtg  cagggaaacg  ggtagttaat  2040
cctgcaagtg  ctttccaggg  attgggttct  aagctgcctc  cagttgttgg  aatagccaat  2100
aagcttaaag  ctaacttaat  gtcagaaacc  aatggtagta  taccctatg  cagtgtattg  2160
ccaaaactga  ttagctagtg  gacagagata  tgccagagag  tgcatatgtg  tatattttta  2220
tgatcttaca  aacctgcaaa  tggaaaaaat  gcaagcccat  tggtggatgt  ttttgttaga  2280
gtagactttt  tttaaacaag  acaaaacatt  tttatatgat  tataaaa  2327

```

&lt;210&gt; 8

&lt;211&gt; 568

&lt;212&gt; PRT

&lt;213&gt; HOMO SAPIENS

&lt;400&gt; 8

```

Met Lys Trp Lys Glu Lys Leu Gly Asp Gly Val Tyr Asp Thr Phe Met
 1          5          10          15
Met Ile Asp Glu Thr Lys Cys Pro Pro Cys Ser Asn Val Leu Cys Asn
          20          25          30
Pro Ser Glu Pro Pro Pro Pro Arg Arg Leu Asn Met Thr Thr Glu Gln
          35          40          45
Phe Thr Gly Asp His Thr Gln His Phe Leu Asp Gly Gly Glu Met Lys
          50          55          60
Val Glu Gln Leu Phe Gln Glu Phe Gly Asn Arg Lys Ser Asn Thr Ile
          65          70          75          80
Gln Ser Asp Gly Ile Ser Asp Ser Glu Lys Cys Ser Pro Thr Val Ser
          85          90          95
Gln Gly Lys Ser Ser Asp Cys Leu Asn Thr Val Lys Ser Asn Ser Ser
          100          105          110

```

Ser Lys Ala Pro Lys Val Val Pro Leu Thr Pro Glu Gln Ala Leu Lys  
 115 120 125  
 Gln Tyr Lys His His Leu Thr Ala Tyr Glu Lys Leu Glu Ile Ile Asn  
 130 135 140  
 Tyr Pro Glu Ile Tyr Phe Val Gly Pro Asn Ala Lys Lys Arg His Gly  
 145 150 155 160  
 Val Ile Gly Gly Pro Asn Asn Gly Gly Tyr Asp Asp Ala Asp Gly Ala  
 165 170 175  
 Tyr Ile His Val Pro Arg Asp His Leu Ala Tyr Arg Tyr Glu Val Leu  
 180 185 190  
 Lys Ile Ile Gly Lys Gly Ser Phe Gly Gln Val Ala Arg Val Tyr Asp  
 195 200 205  
 His Lys Leu Arg Gln Tyr Val Ala Leu Lys Met Val Arg Asn Glu Lys  
 210 215 220  
 Arg Phe His Arg Gln Ala Glu Glu Ile Arg Ile Leu Glu His Leu  
 225 230 235 240  
 Lys Lys Gln Asp Lys Thr Gly Ser Met Asn Val Ile His Met Leu Glu  
 245 250 255  
 Ser Phe Thr Phe Arg Asn His Val Cys Met Ala Phe Glu Leu Leu Ser  
 260 265 270  
 Ile Asp Leu Tyr Glu Leu Ile Lys Lys Asn Lys Phe Gln Gly Phe Ser  
 275 280 285  
 Val Gln Leu Val Arg Lys Phe Ala Gln Ser Ile Leu Gln Ser Leu Asp  
 290 295 300  
 Ala Leu His Lys Asn Lys Ile Ile His Cys Asp Leu Lys Pro Glu Asn  
 305 310 315 320  
 Ile Leu Leu Lys His His Gly Arg Ser Ser Thr Lys Val Ile Asp Phe  
 325 330 335  
 Gly Ser Ser Cys Phe Glu Tyr Gln Lys Leu Tyr Thr Tyr Ile Gln Ser  
 340 345 350  
 Arg Phe Tyr Arg Ala Pro Glu Ile Leu Gly Ser Arg Tyr Ser Thr  
 355 360 365  
 Pro Ile Asp Ile Trp Ser Phe Gly Cys Ile Leu Ala Glu Leu Leu Thr  
 370 375 380  
 Gly Gln Pro Leu Phe Pro Gly Glu Asp Glu Gly Asp Gln Leu Ala Ser  
 385 390 395 400  
 Met Met Glu Leu Leu Gly Met Pro Pro Pro Lys Leu Leu Glu Gln Ser  
 405 410 415  
 Lys Arg Ala Lys Tyr Phe Ile Asn Ser Lys Gly Ile Pro Arg Tyr Cys  
 420 425 430  
 Ser Val Thr Thr Gln Ala Asp Gly Arg Val Val Leu Val Gly Gly Arg  
 435 440 445  
 Ser Arg Arg Gly Lys Lys Arg Gly Pro Pro Gly Ser Lys Asp Trp Gly  
 450 455 460  
 Thr Ala Leu Lys Gly Cys Asp Asp Tyr Leu Phe Ile Glu Phe Leu Lys  
 465 470 475 480  
 Arg Cys Leu His Trp Asp Pro Ser Ala Arg Leu Thr Pro Ala Gln Ala  
 485 490 495  
 Leu Arg His Pro Trp Ile Ser Lys Ser Val Pro Arg Pro Leu Thr Thr  
 500 505 510  
 Ile Asp Lys Val Ser Gly Lys Arg Val Val Asn Pro Ala Ser Ala Phe  
 515 520 525  
 Gln Gly Leu Gly Ser Lys Leu Pro Pro Val Val Gly Ile Ala Asn Lys  
 530 535 540

11

Leu Lys Ala Asn Leu Met Ser Glu Thr Asn Gly Ser Ile Pro Leu Cys  
 545 550 555 560  
 Ser Val Leu Pro Lys Leu Ile Ser  
 565

<210> 9  
 <211> 15  
 <212> PRT  
 <213> HOMO SAPIENS

<400> 9  
 Leu Gly Gly Arg Asp Ser Arg Ser Gly Ser Pro Met Ala Arg Arg  
 1 5 10 15

<210> 10  
 <211> 16  
 <212> PRT  
 <213> HOMO SAPIENS

<400> 10  
 Leu Gly Gly Arg Asp Ser Arg Ser Gly Ser Ala Met Ala Arg Arg Pro  
 1 5 10 15

<210> 11  
 <211> 16  
 <212> PRT  
 <213> HOMO SAPIENS

<400> 11  
 Leu Gly Gly Arg Asp Ser Arg Ser Gly Thr Pro Met Ala Arg Arg Pro  
 1 5 10 15

<210> 12  
 <211> 16  
 <212> PRT  
 <213> HOMO SAPIENS

<400> 12  
 Leu Gly Gly Arg Asp Ser Arg Ser Gly Ser Pro Pro Ala Arg Arg Pro  
 1 5 10 15

<210> 13  
 <211> 16  
 <212> PRT  
 <213> HOMO SAPIENS

<400> 13  
 Leu Gly Gly Arg Asp Ser Arg Ser Pro Ser Pro Met Ala Arg Arg Pro  
 1 5 10 15

<210> 14  
 <211> 16  
 <212> PRT  
 <213> HOMO SAPIENS

12

<400> 14  
Leu Gly Gly Arg Asp Ser Arg Ser Pro Ser Ala Met Ala Arg Arg Pro  
1 5 10 15

<210> 15  
<211> 15  
<212> PRT  
<213> HOMO SAPIENS

<400> 15  
Leu Gly Gly Arg Asp Ser Arg Ala Gly Ser Pro Met Ala Arg Arg  
1 5 10 15

<210> 16  
<211> 16  
<212> PRT  
<213> HOMO SAPIENS

<400> 16  
Leu Gly Gly Arg Asp Ala Arg Ser Gly Ser Pro Met Ala Arg Arg Pro  
1 5 10 15

<210> 17  
<211> 15  
<212> PRT  
<213> HOMO SAPIENS

<400> 17  
Leu Gly Gly Arg Asp Ala Arg Ala Gly Ser Pro Met Ala Arg Arg  
1 5 10 15

<210> 18  
<211> 14  
<212> PRT  
<213> HOMO SAPIENS

<400> 18  
Asn Ile Val Thr Pro Arg Thr Pro Pro Pro Ser Gln Gly Lys  
1 5 10



## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US00/30542

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12Q 1/48; C12N 9/12, 15/54

US CL : 435/15, 194; 536/23.2

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/15, 194; 536/23.2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 860 506 A1 (SMITHKLINE BEECHAM CORPORATION) 26 August 1998, see particularly pages 9 and 10.	1-3
X	US 5,965,420 A (CREASY ET AL.) 12 October 1999, see columns 13 and 14.	1-3
X	GARRET, S. et al. The <i>Saccharomyces cerevisiae</i> YAK1 Gene Encodes a Protein Kinase That is Induced by Arrest Early in the Cell Cycle. Mol. Cell. Biol. August 1991. Vol. 11, No. 8, pages 4045-4052, see entire document.	1-3

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*B* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search  
29 JANUARY 2001Date of mailing of the international search report  
19 MAR 2001Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

*Rebecca Prouty*  
REBECCA PROUTY

Telephone No. (703) 308-0196

BLANK PAGE